Sorting Within and Across French Production Hierarchies

Grigorios Spanos

Working Papers / Documents de travail

Sorting Within and Across French Production Hierarchies

Grigorios Spanos

WP 2015 - Nr 24

Sorting Within and Across French Production Hierarchies*

Grigorios Spanos†

June 15, 2015

Abstract

The objective of this paper is to examine the assignment of workers to layers and firms. In particular, I use an administrative dataset of French workers to study the organization of firms. First, I test whether higher ability workers are employed in the higher layers of firms. Second, I test whether there is positive assortative matching between workers in the different layers of firms. Third, I test whether higher ability workers allow their managers to increase their span of control and employ larger teams. To do this, I first classify employees as residing in different organizational layers such as production and administrative workers, supervisors, senior managers, and owners and CEOs, using occupational codes. From a panel wage regression I then obtain estimates of workers' ability as in Abowd et al. (1999). I then study how workers sort into layers and across layers with other workers. I emphasize three results. First, higher ability workers are employed in the higher layers of firms. Second, I find evidence of positive assortative matching between workers in the different layers of firms. Third, I find different mechanisms are behind the sorting pattern observed in the data. I find evidence that higher ability workers limit their managers' span of control, and I also find weak evidence that higher ability workers allow their managers to increase their span of control and employ larger teams.

KEYWORDS: positive assortative matching, firm organization, matched employer-employee data, high-dimensional fixed effects.

*I wish to thank Kunal Dasgupta, Gilles Duranton and Peter Morrow for excellent supervision. I am indebted to Pierre-Philippe Combes for helping me access and use the data. This paper also benefited from insightful discussions with Bernardo Blum, April Franco, Luis Garicano, Camille Hémet, Francis Kramarz, Nicholas Li, Matthieu Solignac, Marc Sangnier, Michel Serafinelli, Brian Tavares, Daniel Trefler, Ronald Wolthoff and Tanguy van Ypersele, and seminar participants at Brock University, the International Trade Workshop at the University of Toronto, and the AMSE Eco Lunch Seminar at Marseilles. Finally, I would like to thank the staff of the French National Statistical Institute, INSEE, in Paris, for their kindness and support. This project was conducted during my stay at the Aix-Marseilles School of Economics, and funded by the AMSE Pre-Doctoral Research Grant. All remaining errors are my own.

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS.
E-mail: grigorios.spanos@utoronto.ca

1 Introduction

Ever since Coase (1937) economists have known that one of the most important problems a firm faces is how to organize inputs efficiently. However, classical economic models often abstract from firms' organizational decisions. A firm is like a black box, whereby inputs are directly mapped into a final good. However, an understanding of how firms organize is essential, because firms determine the allocation of productive resources in the economy.[1] One important organizational decision of firms is what types and how many workers they should hire, as well as what tasks should be assigned to which workers (Caliendo et al. (2014)).

Despite much theoretical interest, very little is known empirically about how workers sort together in firms.[2] Several researchers have investigated whether good workers are employed in productive firms. While most empirical studies are concerned with how workers match with firms, far fewer studies have examined the different tasks workers perform, and whether better workers are employed with better workers in the other positions of a firm.

This paper fills the gap by examining how workers sort together in firms. My empirical strategy relies on the idea that firms can be thought of as hierarchical teams, composed of layers that perform different tasks. The lowest layer of a firm, for example, contains workers who focus on production, while higher layers contain individuals that perform managerial tasks. With this in mind, I examine how workers sort into teams and layers within each team. More precisely, within a team I first test whether higher ability workers are employed in higher layers. Second, across teams, I test whether there is positive assortative matching, in which the ability of an individual in one layer is positively correlated with the ability of a worker in another layer. Third, I investigate whether this sorting pattern is caused by higher ability workers allowing their managers to increase their span of control and employ larger teams, as suggested by Antras et al. (2006).

I use an administrative dataset of French workers, the Déclarations Annuelles des Données Sociales (DADS) to test these predictions. I begin by classifying employees as residing in the different organizational layers of firms. With my dataset, I observe four distinct layers, production and administrative workers, supervisors, senior managers, and owners and CEOs, by using occupational codes. The concept of a layer that I use is from the management hierarchy theory of the firm that was introduced by Garicano (2000) and used empirically by Caliendo et al. (2014). In theory a layer corresponds to a set employees who earn similar wages, are of similar ability and perform tasks at a similar level of authority (Caliendo et al. (2014)). Since firms are hierarchical teams, layers have the added property that, within a firm, higher layers contain fewer workers

[1] As noted by Rosen (1982): 'The firm cannot be analyzed in isolation from other production units in the economy. Rather, each person must be placed in his proper niche, and the marriage of personnel to positions and to firms must be addressed directly.'

[2] For example, several studies have used models of firm organization to investigate earnings inequality (Garicano and Rossi-Hansberg (2006)), offshoring (Antras et al. (2006),Antras et al. (2008)) and knowledge diffusion (Dasgupta (2012)).

who are of greater ability.

For every firm in the dataset, I calculate the total number of layers in the firm, and the size of each layer, in terms of labor hours worked. With my dataset I can observe four different types of organizations, one-layer firms, two-layer firms, three-layer firms, and four-layer firms. I show that this classification of employees into layers is meaningful and consistent with the concept of a layer discussed above.

Then for the years 1993 to 2004, I use the panel dimension of my dataset to obtain estimates of workers' ability. I estimate a Mincerian wage regression with individual fixed effects, as in Abowd et al. (1999). I use the individual fixed effects from my regression as my measure of worker ability.

Using these measures of the size and number of layers of firms, along with measures of worker ability in these layers, I test my main predictions. First, I conclude that higher ability workers are employed in the higher layers of firms. For example for four-layer firms, I find that an individual with a one hundred percent increase in his ability will on average reside 0.511 layers higher. Second, I find evidence of positive assortative matching between workers in the different layers of firms. For example, in four-layer firms, a one hundred percent increase in the average ability of workers in layer one is associated with a 0.320 increase in the average ability of workers in layer two. Third, I find evidence that higher ability workers limit their managers' span of control. For example, in three-layer firms a one unit increase in the average ability of workers in layer two is associated with a 27.1 percent decrease in their managers' span of control. Finally, I find only weak evidence that higher ability workers allow their managers to increase their span of control and supervise larger teams. For example, in four-layer firms, a one unit increase in the average ability of workers in layer three is associated with a 33.4 percent increase in their managers' span of control.

Together the last two results imply not all firms are organizing their production in a similar manner and that different mechanisms are behind the sorting observed in the data. As long as there is positive assortative matching between workers in the different layers of firms, it is incompatible for higher ability managers to supervise less workers and for higher ability workers to be employed in larger teams. To examine this further I group industries along two different dimensions and re-test my main predictions. I first use the OECD (2003) classification of technological intensity to organize industries by their degree of technological intensity into four categories, low, medium-low, medium-high and high, and second, I use the Rauch (1999) classification of goods and group industries by their degree of product differentiation. In both cases, within groups, I continue to find evidence that high ability workers occupy the upper layers of firms and of positive assortative matching between workers in the different layers of firms. In addition, in medium-low technology industries I find weak evidence that higher ability workers allow managers to supervise larger teams, while in high technology industries I find evidence that higher ability workers limit their managers' span of control. For example in three-layer firms, in medium-low technology firms a unit a one unit increase in the average ability of workers in layer two is associated with a 25.5 percent increase in the span of control of workers in layer three, while in high technology firms it is associated with a 49.6 percent decrease.

Grouping industries by their degree of product differentiation, I find a similar pattern in the data. In industries with a low degree of product differentiation I find weak evidence that higher ability workers allow their managers' to increase their span of control, while in industries with a high degree of product differentiation I find that higher ability workers restrict the number of agents their managers' can supervise. These results suggest that the different mechanisms behind the sorting pattern observed in the data depend on firms' technological intensity and the goods they produce.

In the last part of the paper, I address robustness of my results by assessing several potential threats to my empirical strategy. First, one concern with the empirical analysis is that the worker fixed effects are inconsistent. Because the worker fixed effects are incidental parameters from a wage regression, they can only be measured consistently as the number of years an individual is observed in the panel grows large. To resolve this issue, I conduct my analysis on a restricted sample of worker fixed effects for workers that I observe for at least 10 periods. Second, as discussed in Andrews et al. (2008) my measures of workers' ability may be misestimated and any positive correlation between the individual fixed effects is the result of a positive correlation between the estimated error of the individual fixed effects. These threats are important when examining whether higher ability agents occupy the upper layers of firms, and whether there is a positive correlation between workers' ability in the different layers of firms. To address this issue, I conduct my empirical analysis outside of the sample, for the year 2009, and only on the set of workers who have moved to a firm that they have never been employed in before. Taking both potential threats into account, I continue to find that higher ability workers are employed in the higher layers of firms, and that the ability of individuals in one layer of a firm is positively correlated with the ability of workers in another layer.

This paper is related to the broad literature on the theory of the firm allowing for management hierarchies. With the aim of explaining the distributions of firm size and earnings in the economy, a long-standing literature has examined how productive factors are allocated to managers with different abilities (for example Lucas (1977) and Rosen (1982)). To motivate my empirical strategy, I use a model by Garicano and Rossi-Hansberg (2006) in which agents with different cognitive abilities sort into occupations, layers and teams. Regardless of the distribution of knowledge in the economy, the equilibrium displays skill stratification, in the sense that agents with similar levels of cognitive ability sort into the same occupations and layers across firms. Agents with the least amount of knowledge become production workers, while agents with high levels of ability sort into managerial layers which correspond to higher layers of firms. The equilibrium also displays positive assortative matching, in the sense that higher ability managers organize into firms with higher ability subordinates. The mechanism behind this result is the following:

in a given layer, agents of greater ability can solve a greater proportion of problems, and thus render their subordinates more productive. In turn, because they can solve a greater proportion of problems alone, higher ability subordinates require less of their superiors' time. This frees up the latter's time and allows managers to supervise more workers.

This paper is most closely related to Garicano and Hubbard (2005) who examine positive assortative matching between partners and associates in law firms in Texas.[3] Using data on lawyers' school of education and firm of employment, they find that associates are more likely to work at the same firm as partners who went to a similarly ranked school, consistent with positive assortative matching. The nature of their data, however, does not permit them to obtain a measure of workers' ability that varies across individuals who graduated from the same school. In addition their analysis is limited to two-layer firms: partners and associates. My dataset and classification strategy allow me to make progress on this issue, since I can identify up to four layers in firms. Finally, another distinction is that I examine the mechanism that is causing this sorting pattern: higher ability managers supervising larger teams.

More generally, this paper is also related to a large empirical literature examining sorting in labor markets. In particular, this literature is concerned with whether productive workers are matched with productive firms (see for example Abowd et al. (2003), Abowd et al. (2004), Martins (2008), Andrews et al. (2008)). My paper examines a related but different question: how workers sort with other workers in layers and firms.

My paper is most closely related to de Melo (2013), who uses matched employer-employee data from Brazil to examine whether workers of similar ability sort into the same firms. While Lopes de Melo focuses exclusively on whether workers of similar ability sort into the same firm, I also examine how workers sort into the different layers of firms, and the mechanism that is causing this pattern. In addition I make progress on an econometric issue associated with testing whether higher ability workers sort into the same firms. Specifically, a positive correlation between workers' ability may be due to standard estimation error. I address this issue by examining whether my findings hold in the year 2009, for individuals who have moved to a firm that they have never been employed in before.

The paper proceeds as follows: Section 2 presents a brief description of the management hierarchy theory of a firm and its predictions. Section 3 introduces the data. Section 4 discusses the estimation strategy. Section 5 presents the descriptive statistics and summary results, Section 6 tests the model's predictions, and Section 7 presents robustness checks. Finally, section 8 concludes.

[3] Iranzo et al. (2008) investigate whether production and non-production workers are complements or substitutes. Although in their analysis, managers are contained in their non-production worker classification, they do not focus on the relationship between managers and production workers.

2 Model

In this section I briefly present the knowledge-based management hierarchy models from Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and discuss their main implications. The models are concerned with the organization of teams and in this study I interpret teams to be firms. To fix ideas, I first present the model where teams have three layers. For a complete exposition and for all proofs, I refer the reader to Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and Antras et al. (2008).

2.1 Setup of the Model

In the model, a unit of output is produced only when a problem is solved. Problems are differentiated by their difficulty and for simplicity, assume that the difficulty of problems is drawn from a uniform distribution with support $[0, 1]$. Agents have one unit of time and are heterogeneous in their level of knowledge.[4] Agents' knowledge determines their ability to solve problems, and assume that knowledge is cumulative: an individual with knowledge z can solve all problems in the interval $[0, z]$.[5]

Production occurs in teams. Teams are composed of one layer of production workers, who spend their one unit of time drawing problems and attempt to solve them, and layers of managers who do not draw problems but instead spend all of their time solving problems that their production workers cannot solve.[6] For managers to receive problems that other agents cannot solve, communication is possible between managers and production workers within a team. Communication, however, entails a cost to the managers of a team.

In a team, production workers draw one problem per unit of time and if they can solve the problem a unit of output is produced. Otherwise, they ask their manager in the immediate layer above who in turn spends a fraction of her time communicating with the worker. If the manager knows the solution to the problem, then she conveys the solution to her worker who immediately produces a unit of output. If the manager does not know the solution, the production worker asks the manager two layers above. This process continues until the problem is solved, or the production worker has seen a manager in every layer of the team, at which point the problem remains unsolved and nothing is produced.

Consider a team composed of three layers, one manager in layer three with knowledge z_m^3, n_m^2 managers in layer two with knowledge z_m^2 and n_p^1 production workers in layer one with

[4] I abstract from the decision to acquire knowledge. For a model where agents acquire knowledge see Garicano and Rossi-Hansberg (2006).

[5] The output of such an individual from working alone is therefore:

$$y(z) = z. \tag{1}$$

[6] In other words, production workers specialize in routine tasks (i.e. production), while managers specialize in nonrountine tasks (i.e. problem solving).

knowledge z_p^1.[7] Let h denote the time cost per problem that a manager incurs communicating with a production worker and assume that this cost is the same across all managers. The number of managers in layer two, n_m^2, is determined by the number of problems in the team, the fraction of problems production workers cannot solve, and the communication costs. It is equal to:

$$n_p^1 h[1 - z_p^1] = n_m^2, \tag{2}$$

where $h[1 - z_p^1]$ is the total cost per unit of time that a manager in layer two incurs while working in a team composed of production workers with knowledge z_p^1. Similarly, since managers in layer two can solve z_m^2 fraction of problems, the time constraint of a manager in layer three is equal to:

$$n_p^1 h[1 - z_m^2] = 1, \tag{3}$$

where $h[1 - z_m^2]$ is the total cost per unit of time that the top manager incurs while working in a team composed of managers in layer two with knowledge z_m^2. The communication technology therefore limits the amount of interactions managers can have with their subordinates, and this in turn determines the number of production workers, n_p^1, and the number of managers in layer two, n_m^2.

Furthermore, the output of the team is determined by number of problems in the team and the fraction of problems the team solves. The latter is determined by the knowledge of the top manager. Since the manager in layer three receives all problems that her production workers and her managers in layer two cannot solve, and she can only solve a fraction z_m^3 of them, the output of the team is equal to:

$$y(z_m^3, z_m^2, z_p^1) = n_p^1 z_m^3. \tag{4}$$

This production framework has several important properties that determine the equilibrium allocation of managers and production workers to teams. First, regardless of their occupation, agents are not perfect substitutes in production. Second, in this framework matching is many to one and because they share their knowledge with other agents in a team, managers increase the value of their knowledge by concentrating on problems that only they can solve. Third, production is asymmetrically sensitive to skill. Since they can leverage their knowledge over many workers, managers are more important to the output of a team. And fourth, agents in the different layers of a team are complements. The mechanism behind this result is the following. Managers of greater knowledge can solve a greater proportion of problems, and thus render subordinates more productive. In turn, more knowledgeable subordinates increase the productivity of their managers. Since all individuals have one unit of time, managers are constrained in the number of agents that they can supervise and because they can solve a greater proportion of

[7] I slightly depart from the notation in Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and refer to teams by the total number of layers instead of the number of management layers.

problems on their own, more knowledgeable subordinates spend less time communicating with managers which allows the latter to supervise larger teams.

In equilibrium the assignment of agents into occupations and teams has the following properties. First because production is asymmetrically sensitive to skill, there is skill stratification in the sense that agents with greater knowledge sort into managerial occupations while agents with lesser knowledge become production workers. Second because they are complements there is positive assortative matching between agents in the different layers of teams. Third, managers of greater knowledge supervise more subordinates and form larger teams, and fourth subordinates of greater knowledge work in larger teams. These two properties follow from positive assortative matching between agents in the different layers of teams and managers' time constraint.

Furthermore, these results are generalizable to an economy where teams can have any number of layers (see Garicano and Rossi-Hansberg (2006) for details). An important point to note is that any model with a production function that exhibits similar interactions between managers and production workers will, in a general equilibrium, yield similar results (Garicano and Hubbard (2008)).[8] More specifically, as long as high skill agents raise the productivity of their subordinates, and better individuals require less supervision, then in equilibrium high skill individuals will form teams with more and better subordinates in the layers below. These properties are summarized in the proposition below:

Proposition 1 *With L layer teams, the equilibrium assignment of individuals to occupations and teams has the following properties:*

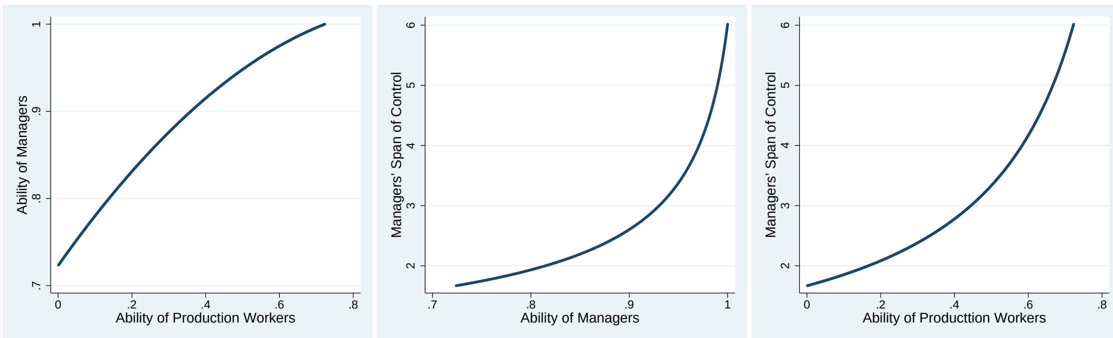
- *Individuals with the greatest knowledge sort into the top managerial occupations, individuals with intermediate knowledge sort into managerial occupations, while individuals with least knowledge become production workers.*
- *There is positive assortative matching between agents in the different layers of teams.*
- *For all layers, agents with greater knowledge are employed in larger teams and are supervised by less individuals in the layers above.*
- *For all layers, managers with greater knowledge are employed in larger teams and supervise more individuals in the layers below.*

2.2 Discussion: Taking the Model to the Data

In this section thus far, I have presented the intuition of the model and its predictions. The sections further below are concerned with testing the model's predictions. The paper's empirical strategy relies on the idea that firms can be thought of as hierarchical teams, composed of layers

[8] Eeckhout and Kircher (2012) provide a general formulation of production in two-layer teams, and conditions under which the equilibrium exhibits positive assortative matching and better managers supervise larger teams.

Figure 1: Equilibrium in Economy with Two-Layer Firms



(a) Assignment of Workers to Managers (b) Span of Control on Managers' Knowledge (c) Span of Control on Subordinates' Knowledge

that perform different tasks. Henceforth I will use teams and firms interchangeably. In addition, in my analysis I relate agents' knowledge in production to their productivity and their ability to produce.[9] Henceforth I will simply refer to agents' knowledge as their ability.

To examine whether abler agents occupy the upper layers of firms I regress estimates of workers' ability on their position within firms. To examine whether there is positive assortative matching between workers in the different layers of firms I build on recent tests suggested in the literature investigating matching between workers and firms (de Melo (2013)), and estimate a correlation between the average ability of workers in the different layers of firms. To test whether abler managers supervise more subordinates, I rely on the equations that characterize managers' time constraint. Equations (2) and (3) can be generalized to a firm with L layers. Rearranging, taking logs, and defining the span of control of workers in layer $l+1$ as the ratio of the size of layer 1 to the size of layer $l+1$ one obtains the following expression:

$$\begin{aligned} \ln span^{l+1} &= \ln \frac{n^1}{n^{l+1}} \\ &= \ln h - \ln[1 - z^l]. \end{aligned} \tag{5}$$

where n^1 denote the size of layer 1, n^{l+1} denotes the size of layer $l+1$ and z^l denotes the ability of agents in layer l. As managers occupy all other layers except layer 1, I drop the subscripts p and m from the notation. These equations have the following characteristics in common: they are a function of the number of production workers, and depend on the knowledge of the individuals in the layer below. According to the model, therefore, managers' span of control should be increasing with their subordinates ability. To test this prediction, I approximate equation (5) with the following equation:

[9] This is not an overly strong assumption. The knowledge-based hierarchy models of Garicano (2000) have a specific formulation, however as mentioned above, any model with a production function that has the same interactions between managers and production workers will yield similar results (Garicano and Hubbard (2008)).

$$\ln span^{l+1} = \gamma_0 + \gamma_1 \widehat{z^l} + u, \tag{6}$$

where $\widehat{z^l}$ represents my measure of the ability of agents in layer l and u is an error term. If the mechanism described by the model, that determines how agents sort together into firms, holds in the data then the estimated coefficient γ_1 should be positive and significant.

Furthermore, since there is positive assortative matching, there is a one-to-one correspondence between the ability of individuals in the different layers of a firm. Therefore, equation (5) can also be rewritten as:

$$\ln span^{l+1} = \ln h - \ln[1 - f(z^{l+1})]. \tag{7}$$

where $f()$ is a function that maps the ability of workers in layer l to layer $l+1$; i.e. $z^l = f(z^{l+1})$. I approximate equation (7) with the following equation:

$$\ln span^{l+1} = \gamma_0 + \gamma_1 \widehat{z_{l+1}} + u, \tag{8}$$

where $\widehat{z_{l+1}}$ represents my measure of the ability of agents in layer $l+1$ and u is an error term. If the mechanism described by the model holds in the data then the estimated coefficient γ_1 should be positive and significant.[10]

Furthermore, the theory describes how agents sort together into firms. An important point to note is the theory presented above is very general. The theory is silent on whether firms can operate many establishments, in many industries or in many locations. It is likely that firms operating in many industries and locations adopt a different organizational structure. Therefore, in the sections below I take a strict interpretation of the model and examine its predictions only on firms that operate in a single industry and a single location.[11] The theory is also silent on whether agents sort into industries and locations before they sort into firms. Indeed, it is likely that the production technology and, in particular, the problems agents encounter are different across industries. In the sections further below, I therefore present results with different sets of controls so as to examine whether the model's predictions hold in the aggregate economy, within sectors, and within locations.

[10] To test for the mechanism I use both approaches because, as discussed below, workers' ability is estimated from data that do not contain all individuals in the economy.

[11] Firms operating in many industries are multi-product firms and their organization is likely to be affected to the number of products that they produce. Also if firms operate in many locations, then it is likely that their organization will have a different structure. See for example the study by Antras et al. (2008) for the case of firms operating in different locations.

3 Data Description & Classification of Layers

The data are extracted from the Déclarations Annuelles des Données Sociales (DADS), which are provided and maintained by the French National Statistical Institute for Statistics and Economic Studies (INSEE). The DADS are matched employer-employee datasets and are constructed from administrative records that must be completed by all employers in France. A report must be filled by each establishment for every one of its employees, so there is a unique record for each employee-establishment-year combination. The DADS contains two datasets: a panel of workers born in October and that runs from 1976 to 2009, and from 1993 to 2009, exhaustive cross-sections of all workers in mainland France.[12]

In both the panel and cross-section datasets, for each observation, there is information on employees' characteristics, such as age, gender, and occupation, basic information on the establishment, such as location, industry and the parent firm, and basic firm level information, such as the firm's industry. For each observation there is also information on annual earnings, denominated in 2007 euros, number of days worked, and number of hours worked.[13]

As discussed further below, I use the panel dataset to obtain measures of workers' ability. For computational tractability, I restrict the sample to the years 1993 to 2004, and to all full-time workers who are born in October in an even numbered year, are between the ages 18 and 65 and work in mainland France. For the years 1993 to 2004, there are 4,999,728 observations, corresponding to 753,092 workers in 399,676 firms. Appendix A and Abowd et al. (1999) provide further details on the data and information on how wages are determined in France.

For the year 2004, I use information from the cross-section to measure the total number of layers in firms and the size of each layer.[14] In the management hierarchy theory of the firm by Garicano (2000) a layer corresponds to a set employees who earn similar wages, are of similar ability and perform tasks at a similar level of authority (Caliendo et al. (2014)).[15] To construct the different layers of firms, I adopt the strategy put forth by Caliendo et al. (2014), and use one-digit occupational codes, which range from 2 to 6, to classify employees into layers.[16] In total, I can classify employees into four distinct layers. Layer 1 corresponds to qualified and non-qualified administrative workers and blue-collar workers. It contains all workers with occupational codes 5 and 6. Layer 2 is composed of supervisors and individuals with a higher level of responsi-

[12]Until 1993 the DADS only contained information on individuals born in October in an even numbered year. From 1993 onwards, the DADS contains information on all individuals born in October.

[13]Information on the total number of hours worked is only available after 1993.

[14]Since the panel of the DADS is only a 5 percent sample of the population, it is not suitable to properly measure the total number of layers and the size of each layer in a firm. Appendix A and Caliendo et al. (2014) provide further details on the exhaustive cross-section data.

[15]The concept of a layer that I use is therefore independent of the actual occupations of employees, such as whether they are lawyers, engineers or computer programmers. Instead it depends on their knowledge, productive ability, and their relative position in the organizational hierarchy of firms. In addition, since firms are hierarchical, layers have the added property that within a firm higher layers contain less workers who are of greater ability.

[16]The occupational codes range from 1 to 6. I have removed all firms operating in the agricultural and fishing industries, which correspond to occupational code 1.

Table 1: Description of Firms by Total Number of Layers

Total Number of Layers	Average Number of employees	Average Number of hours	Average Wage	Average Ability	Median Wage	Median Ability	Standard Deviation Wage	Standard Deviation Ability
ONE								
1st layer	11.00	11,844.12	2.13	0.433	2.10	0.442	0.195	0.435
TWO								
1st layer	17.16	21,874.82	2.18	0.347	2.14	0.358	0.231	0.411
2nd layer	2.71	4,115.16	2.40	0.428	2.38	0.419	0.288	0.409
THREE								
1st layer	49.16	71,431.91	2.33	0.352	2.27	0.357	0.321	0.375
2nd layer	11.72	19,900.35	2.61	0.473	2.57	0.496	0.352	0.365
3rd layer	6.25	10,350.94	2.96	0.580	2.96	0.594	0.349	0.402
FOUR								
1st layer	58.57	85,769.69	2.36	0.359	2.29	0.365	0.335	0.381
2nd layer	14.90	25,643.13	2.64	0.501	2.61	0.519	0.361	0.372
3rd layer	8.91	14,515.65	3.00	0.574	2.98	0.582	0.362	0.389
4th layer	1.22	2,165.00	3.27	0.772	3.39	0.765	0.378	0.398

Notes: Descriptive statistics of firms that are in both the exhaustive cross-section and panel datasets of the DADS, for the year 2004. These statistics are reported separately for firms with different number of layers. Column 1 refers to the layer within a firm. Columns 2 report the average number of employees in a given layer, while column 3 reports the average number of hours worked by employees in a given layer. These measures are obtained from the exhaustive cross-section of the DADS. Columns 4, 6 and 7 report the average log-hourly wages, median log-hourly wages and standard deviation of log-hourly wages within a layer. Columns 5, 7 and 9 report the average ability, median ability and standard deviation of ability of workers in a given layer. Measures of wages and ability values are obtained from the panel dataset of the DADS. Ability is estimated from equation (9).

bility than ordinary workers, and contains all workers with an occupational code 4. Layer 3 is composed of senior directors and top management staff and contains all workers with an occupational code 3. And, layer 4 corresponds to owners who receive a wage and CEOs. It contains all workers with occupational code 2.

Firms can have as many as four layers in their organization. I consider a firm to have a layer if there is at least one employee in the exhaustive cross-section employed there and retain only firms with at least one employee in layer 1.[17] Finally, I merge the information from both the panel and exhaustive cross-section datasets together, and retain only firms that operate in non-service sectors, and firms that operate in only one industry and location.[18] In all, the matched dataset contains 18,790 firms that operate in 17 industries, of which 2,160 are one-layer firms, 3,322 are two-layer firms, 7,860 are three-layer firms and 5,450 are four-layer firms.[19]

Table 1 reports summary statistics of firms in the matched sample. In the table firms are grouped by their organizational structure, their total number of layers. The average firm in the

[17] In other words, I remove firms that do not have any employees in layer 1. The appendix provides further details on the construction of layers within firms.

[18] Locations correspond to the 341 employment areas in mainland France.

[19] Appendix A provides further information on the number of firms and workers by industry in the data. Unlike the exhaustive cross-section, since the panel data is based on a 5 percent sample of the French population, it contains information on a sample of firms operating in mainland France.

sample has an organizational structure that is consistent with the knowledge-based management hierarchical theory of a firm. On average lower layers of firms are larger than the layers above, and contain workers that earn lower wages and are of lower ability. For example for the average four-layer firm, layer 1 has 58.57 employees, layer 2 has 14.90 employees, layer 3 has 8.91 employees and layer 4 has 1.22 employees. The findings are similar if one measures the size of layers by the number of hours worked. Returning again to four-layer firms, the average log-hourly wages of workers in layer 4 are 3.27, the average log-hourly wages of workers in layer 3 are 3.00, and the mean log-hourly wages of workers in layers 2 and 1 are 2.64 and 2.36 respectively. Therefore there is a clear ranking in wages. The same ranking also holds for ability, where for example in four-layer firms workers who reside in layer 4 have the greatest average ability, and are succeeded by layers 3, 2, and 1.[20] The classification of workers into layers, therefore has economic meaning. The evidence is consistent with the view that firms are hierarchies, in the sense that higher layers of a firm are smaller and contain workers earning higher wages and of greater ability.

4 Estimating Ability

To obtain measures of workers' ability from the data, I use the empirical approach of Abowd et al. (1999) which has been developed further by Card et al. (2013). I model log-hourly wages, w_{it}, of worker i in time t, as a linear function of a time-invariant worker component θ_i, a time-invariant firm-layer component $\psi_{J(i,l,t)}$, time varying worker characteristics x_{it}, and a mean-zero error term ϵ_{it}. The equation to be estimated is:

$$\ln w_{it} = x_{it}\beta + \theta_i + \psi_{J(i,l,t)} + \epsilon_{it}. \tag{9}$$

The term θ_i captures the portable part of a worker's wages that remain with him as he moves across firms, or layers within firms. The variation of this term reflects a worker's productivity, bargaining ability and labor market discrimination (Card et al. (2013)). In the subsequent analysis, I use θ_i as my measure of workers' ability. The terms x_{it} captures how workers' earnings evolve with changes in their observable attributes, such as labor market experience. In my estimation, I use age as a proxy for experience and the list of time-varying controls in x_{it} are age and age squared, both interacted with gender.[21] Although in theory, workers only form firms with other workers, I include in equation (9) firm-layer fixed effects, $\psi_{J(i,l,t)}$, which are meant to identify firm attributes that affect every worker's earnings in a given layer in a firm equally, such as compensation policies, bargaining strength in the labor market, and productivity. Alternatively, since not all workers are employed in the same firm throughout their career, one can interpret the firm-layer fixed effects as partially accounting for any permanent influences past employees

[20] The following section explains how I estimate workers' ability.

[21] Since in equation (9) age cannot be separately identified from worker and time fixed effects, I exclude any time fixed effects from the analysis. Indeed one can show that age can be written as a linear combination of the time and worker fixed effects.

may have on the current organization, or any influences that affect individuals' earnings in a given layer of a firm that are the result of workers in the other layers.

To identify all of the econometric parameters in equation (9), I assume, as in Abowd et al. (1999), that the error term ϵ_{it} is strictly exogenous. Under this assumption, the parameter β can be consistently estimated as the number of workers, N, the number of firm-layers, J, and the number of years, T, increases. The parameters θ_i and $\psi_{J(i,l,t)}$ can only be separately identified by workers who switch employers, or layers within employers in the panel. In the dataset, there are in total of 1,156,816 worker displacements. Since θ_i is an incidental parameter, consistent estimates for it can only be obtained as the number of years a worker is observed grows large. Table 1 in the Appendix A presents the distribution of the number of years a worker in observed in the panel. Over 50 percent of workers are observed for 6 years or more. Similarly, $\psi_{J(i,l,t)}$ can only be consistently estimated if the number of workers in a layer in a firm, or the number of years grows large. Table 2 in the Appendix A presents the distribution of the number of workers observed in a layer in a firm in a given year, as well as the number of years firms' layers are observed in the panel. The average number of workers in a layer in a firm is 2.67, and over 50 percent of firms' layers are observed for 2 years or more.

To estimate equation (9) I focus on the largest connected group, that is the largest group of layers within firms that, over the years, have had at least one employee in common with another layer in the same or a different organization. In the panel, the largest connected group contains 753,092 workers and 569,198 layers within firms. To estimate equation (9), I use the algorithm put forth by Guimaraes and Portugal (2010), which builds on the algorithm of Abowd et al. (2003).[22]

5 Results & Descriptive Statistics

Table 2 summarizes the estimation results from regression (9).[23] To summarize my findings I report the standard deviation of log-hourly wages, of the worker and firm-layer effects. I also report the root mean squared error (RMSE) of the residuals and the adjusted R-squared of the estimation. One important point to note is that the standard deviation of the worker effects is less than the standard deviation of wages. In the model because workers of different abilities are more productive from working in firms rather than alone, individuals' wages are amplified relative to their ability, and hence the standard deviation of wages is greater than the standard deviation of abilities, consistent with the data.

In Table 2, I also report correlations. The correlation between the worker and firm-layer fixed effects is -0.1636. This finding is similar to the empirical literature that investigates how workers sort into firms. As many researchers report there is a negative correlation between worker and

[22]The solution of the algorithm provides a non-unique set of solutions for the worker and firm-layer fixed effects. To render the effects unique the algorithm sets the average of the firm-layer fixed effects to zero.

[23]This table has a similar structure to Table III in Card et al. (2013).

Table 2: Summary Statistics

Sample Year	$1993-2004$
Worker and Firm-Layer Parameters	
Number of Worker Effects	$753,092$
Number of Firm-Layer Effects	$569,198$
Summary of Parameter Estimates	
St. Dev. of Wages	0.4417
St. Dev. of Worker Effects	0.3940
St. Dev. of Firm-Layer Effects	0.2508
RMSE of AKM Residual	0.1717
Adjusted R-Squared	0.8489
Correlations	
Wages & Worker Effects	0.2509
Wages & Firm-Layer Effects	0.5073
Worker Effects & Firm-Layer Effects	-0.1636
Comparison with the Match Effects Model	
RMSE of Match Model	0.1490
Adjusted R-Squared	0.8862
ADDENDUM	
Sample Size	$4,999,728$

Notes: Results from OLS estimation of equation (9). $X\beta$ includes age and age squared interacted with gender. The match model includes $X\beta$ and a separate dummy for each worker-firm pair, corresponding to a job.

firm fixed effects, estimated from a log-linear wage equation.[24] In my analysis, I abstract from this correlation, since I am concerned with how employees in each layer of a firm match, rather than how workers match with firms.[25] Furthermore, the correlation between the individual fixed effects and log-hourly wages is 0.2509. Therefore, individuals of higher ability earn more.

Table 2 also contains the adjusted R-squared and RMSE of a model with unrestricted match effects, that is a separate dummy for each worker-firm-layer job spell. If match effects are an important determinant of workers' wages, then a model with worker-firm-layer match effects should provide a markedly better fit to the data. The match effects model has an adjusted R-squared of 0.8862 and a RMSE of 0.1490, while the adjusted R-squared from the estimation of equation (9) is 0.8489 and the RMSE is 0.1717. The match effects model, therefore, fits the data slightly better than a specification with separate worker and firm-layer effects. Although this indicates that a match component is present in wages, the improvement in fit is modest.

As in Card et al. (2013), I further examine the importance of a match component to wages. In particular, I examine the wage dynamics of all individuals who changed firms, or layers within firms, in the years 1993 to 2004 with at least two consecutive years in the new and old position. I

[24] See for example the studies Abowd et al. (2003), Abowd et al. (2004), Martins (2008), Andrews et al. (2008).

[25] The approach in this paper is therefore consistent with the papers of de Melo (2013) and Eeckhout and Kircher (2011).

classify the origin and destination positions by the quartile of the estimated firm-layer effects and calculate the average hourly wages of agents in each cell two years before and after the move. I report the results in Table A1. If the error term in equation (9) is exogenous, changes in the wages of individuals who transition from one quartile to the other should be relatively symmetric, and individuals who move to new firms, or layer within firms, within the same quartile should not experience a wage gain. In addition, the increase in wages of workers who transition to different quartiles should be monotonically increasing with the distance of the quartiles.[26] These conditions hold in Table A1. As a visual aide Figure 2 panel (a) illustrates the wage profiles of workers in the first and fourth quartiles. The gains or losses of individuals who transition to quartiles is monotonically increasing with the distance between the quartiles, and the gains or losses are relatively symmetric. Panel (b) illustrates the wage profiles of workers that remain within the same quartile. These profiles are relatively flat. Therefore, at a minimum, the model in equation (9) is a relatively decent first approximation to wages.

6 Tests

6.1 Examining Sorting into Layers: Skill Stratification

I first test for skill stratification, that abler individuals occupy the upper layers of firms.[27] For individual i employed in firm $j(i)$, I estimate the following equation:

$$layer^{L}_{j(i)} = \mu_0 + \mu_1 ability_i + X_{j(i)} + u_{j(i)}, \tag{10}$$

where $layer^{L}_{j(i)}$ is the layer in firm j with a total number of L layers that worker i occupies, $ability_i$ is the estimated ability of worker i, and $X_{j(i)}$ are industry and location controls.[28] Equation (10) is estimated across firms with the same total number of layers, separately. In equation (10) the interest is in how agents sort into layers and if abler individuals occupy the upper layers of organizations then μ_1 will be positive and significant.

Table 3 reports regression results. Each entry in the table reports the estimated coefficient of μ_1. Because of the large number of indicator variables in the regressions, I estimate equation (10)

[26] Consider a worker employed in layer 1 in Firm 1 in period t who moves to layer 1 in Firm 2 in period $t+1$. Ignoring the returns to observables, his expected change in wages is equal to: $E[w_{it+1} - w_{it}|J(i,l,t+1) = \{1,2\}, J(i,l,t) = \{1,1\}] = \psi_{1,2} - \psi_{1,1} + E[\epsilon_{it+1} - \epsilon_{it}|J(i,l,t+1) = \{1,2\}, J(i,l,t) = \{1,1\}]$. Similarly, the expected change in wages of a worker moving in the other direction is equal to: $E[w_{it+1} - w_{it}|J(i,l,t+1) = \{1,1\}, J(i,l,t) = \{1,2\}] = \psi_{1,1} - \psi_{1,2} + E[\epsilon_{it+1} - \epsilon_{it}|J(i,l,t+1) = \{1,1\}, J(i,l,t) = \{1,2\}]$. Therefore if the error term is strictly exogenous both expressions are simply equal to $\psi_{1,2} - \psi_{1,1}$ and $\psi_{1,1} - \psi_{1,2}$. See Card et al. (2013) for further explanations.

[27] Examining whether this sorting pattern is present in the data is important because it confirms that workers in a layer are abler than their subordinates. If this were not the case then the model's prediction are false. Further the argument made in the model is that organizations exist to optimally utilize the knowledge of their workers. By shielding knowledgeable agents from easy tasks, a hierarchy allows abler individuals to focus on solving more complex or harder problems, while lower ability individuals focus on easier or commoner problems. This implies that within a firm, higher ability agents occupy the upper layers of organizations.

[28] There are 17 industries in the dataset and locations correspond to the 341 employment areas in mainland France.

Table 3: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	(1)	(2)	(3)	(4)	(5)
TWO	4,432	0.053***	0.044***	0.047***	0.043***	0.041
		(0.013)	(0.012)	(0.014)	(0.013)	(0.035)
THREE	19,841	0.369***	0.347***	0.349***	0.336***	0.411***
		(0.015)	(0.015)	(0.014)	(0.014)	(0.017)
FOUR	16,003	0.469***	0.446***	0.438***	0.428***	0.511***
		(0.018)	(0.017)	(0.016)	(0.017)	(0.020)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 iterations) in parentheses. OLS regressions for equation (10). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France.*

using OLS. Rows 1 to 3 contain the results for firms with 2, 3 and 4 layers in their organization, respectively. The first column in Table 3 indicates the total number of layers in firms, the second column contains the sample size of the regressions, and the third to seventh columns report the estimated value of the coefficient, μ_1. All the standard errors are bootstrapped and based on 500 replications.

The regressions in column (1) report how agents sort into layers across firms, industries and locations. In all three regressions the coefficients are positive and significant at the one percent level. The column indicates that an individual with a one hundred percent increase in his ability and employed in a two-layer firm will on average reside 0.053 layers higher, while if he is employed in an organization with three-layers he will on average reside 0.369 layers higher and if he is employed in a four-layer firm he will on average reside 0.469 layers higher.

Even within industries and locations, Table 3 reports that higher ability agents occupy the upper layers of firms. To examine how agents sort into layers within industries, in column (2) I include industry fixed effects. The coefficients remain positive and significant at the one percent level. In column (3) I include location fixed effects so as to examine how agents sort into layers within locations. The coefficients in column (3) remain positive and significant. And finally, in column (4) I include both industry and location fixed effects. The findings indicate that an individual with a one hundred percent increase in his ability and employed in a two-layer firm will on average reside 0.043 layers higher, if he is employed in a three-layer firm will on average reside 0.336 layers higher, and if he is employed in an organization with four layers he will on average reside 0.428 layers higher. Therefore even within industries and within locations abler individuals occupy the upper layers of organizations.

The regressions in column (5) contain firm fixed effects and examine how agents sort into layers within firms. In two out of three regressions, the coefficient μ_1 is positive and significant at the one percent level. For four-layer firms, the result indicates that an individual with a one hundred percent increase in his ability will on average reside 0.511 layers higher. For organizations with three layers an agent with one hundred percent increase in his ability will on average reside 0.411 layers higher. Within three and four-layer firms, therefore, abler individuals are employed in the upper layers of organizations.

In column (5), the value of μ_1 for two-layer firms is 0.041 and is not significant at the ten percent level. This would suggest that within two-layer firms, higher ability agents are not sorting into the upper layers of organizations. However, in light of the fact that the coefficient of μ_1 in column (5) is similar in magnitude to column (4), and that in the average two-layer firm there are 1.3 observations in the dataset, these findings are inconclusive.

To summarize from the evidence presented in Table 3 one can conclude that there is skill stratification, in the sense that abler agents occupy the higher layers of organizations. In other words, agents in higher layers of firms are of greater ability than their subordinates in the layers below. I now proceed to examine whether there is positive assortative matching between agents in the different layers of firms, and whether the mechanism that determines the sorting pattern is as suggested by the model.

6.2 Examining Sorting into Teams: Positive Assortative Matching

I now test for positive assortative matching between workers in the different layers of firms.[29] According to the knowledge-based management theory of firms layers are composed of workers who are of similar ability. A representative measure of a layer's ability is the weighted average ability of workers occupying a layer.[30] More specifically, let N_j^l and H_j^l denote the number of individuals and the total number of hours in layer l at firm j, and $H_{j(i)}^l$ denote the number of hours performed by individual i in layer l in firm j. A representative measure of the ability of all workers in layer l at firm j is the following:

$$ability_j^l = \sum_{i=1}^{N_j^l} \frac{H_{j(i)}^l}{H_j^l} ability_i. \tag{11}$$

where the summation is taken over all individuals in layer l at firm j.[31]

[29]Appendix B contains additional tests for positive assortative matching. Appendix B tests whether better workers sort into organizations with better co-workers. I adopt the approach of de Melo (2013), and investigate whether a worker's fixed effect is positively correlated with that of his co-workers. I conduct this analysis across several dimensions and find evidence in favor of positive assortative matching.

[30]I use a weighted average to account for the fact that some workers may be employed for the full year in a firm. In such a case, these workers cannot have the same impact on a firm, as workers who have been employed for the entire year.

[31]Note that this construction is only possible for layers where I there is at least one employee in the panel dataset of the DADS. As the panel is only a five percent sample of the French population, for many firms $ability_j^l$ remains

In firms with the same organization, positive assortative matching implies there should be a positive correlation between the ability of workers in the different layers of firms. For example, when comparing two two-layer firms, the firm with the better production workers in layer 1 will also employ the better managers in layer 2. To test for positive assortative matching I therefore estimate the following equation:

$$ability_j^l = \alpha_0 + \alpha_1 ability_j^{l-g} + X_j\beta + u_j, \tag{12}$$

where $ability_j^l$ is the estimated weighted average ability of all workers in firm j who are in layer l, and $ability_j^{l-g}$ is the estimated weighted average ability of all workers in firm j who are in layer $l-g$, for $g = 1, .., l-1$. The firm controls X_j are firm observable variables such as firm age, an indicator for whether the firm already existed in the first year I have information, 1976, as well as indicator variables for industry and location. I include industry and location fixed effects because the assignment of workers to layers and firms may be different across industries and locations. I estimate equation (12), for firms with the same organizational structure and for the different values of l and g. In equation (12) the interest is how $ability_j^l$ varies with $ability_j^{l-g}$ across firms with the same total number of layers. If there is a positive assortative matching, then the coefficient α_1 will be positive and significant.

Table 4 reports results. Each entry in the table illustrates the estimated coefficient of α_1 between two layers. The first column indicates the total number of layers in firms. The second column indicates the layer for which weighted average ability is the left-hand-side variable in equation (12), and the third column indicates the layer for which weighted average ability is the right-hand-side variable in equation (12). The fourth column reports the sample size of the regressions, while the fifth to tenth columns report estimated values of the coefficient.[32] In Table 4 all the standard errors are bootstrapped and based on 500 replications.

Column (1) contains no controls and tests how agents sort together into firms across industries and locations. In these regressions, the majority of the estimated values of α_1 are positive and significant, indicating that across industries and locations there is positive assortative matching. For example, in organizations with four layers, a one unit increase in the average ability of workers in layer one is associated with a 0.341 average increase in the average ability of workers in layer two.

Even within industries and locations, there is evidence of positive assortative matching. Column (2) contains industry fixed effects and examines how agents sort together within industries,

undefined.

[32] In Table 4 sample sizes vary across regressions and holding organization fixed, sample sizes increase when equation (12) is estimated with lower layers. For example, in regressions with four-layer firms the sample can be as small as 15 observations or as large as 1,249 observations. The reasons are twofold. First given the nature of the data, I do not observe workers in all layers of firms. And second, in the data I am more likely to observe a worker in the lower layer of an organization. That is if firms are hierarchies then there are more workers in the lower layers of firms. Therefore in a 1/12 random sample of the population of workers one is more likely to observe individuals employed in the lower layers of firms.

Table 4: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	(1)	(2)	(3)	(4)	(5)	(6)
TWO	2	1	142	0.103	0.116	0.004	0.145	0.110	0.155
				(0.092)	(0.097)	(0.163)	(0.228)	(0.092)	(0.302)
THREE	3	2	457	0.240***	0.240***	0.223***	0.224***	0.241***	0.227***
				(0.068)	(0.071)	(0.081)	(0.083)	(0.069)	(0.083)
THREE	3	1	662	0.302***	0.315***	0.279***	0.297***	0.303***	0.292***
				(0.064)	(0.064)	(0.078)	(0.081)	(0.063)	(0.081)
THREE	2	1	1385	0.233***	0.217***	0.250***	0.235***	0.235***	0.237***
				(0.033)	(0.033)	(0.037)	(0.037)	(0.033)	(0.037)
FOUR	4	3	15	0.636*				0.691	
				(0.370)				(0.479)	
FOUR	4	2	22	−0.181				−0.118	
				(0.290)				(0.282)	
FOUR	4	1	37	0.333	0.363	0.559		0.417	
				(0.300)	(0.256)	(1.189)		(0.322)	
FOUR	3	2	452	0.195***	0.178***	0.215***	0.169*	0.194***	0.167*
				(0.061)	(0.065)	(0.082)	(0.099)	(0.061)	(0.101)
FOUR	3	1	687	0.077	0.072	0.115	0.117	0.075	0.118
				(0.057)	(0.059)	(0.072)	(0.076)	(0.056)	(0.077)
FOUR	2	1	1249	0.341***	0.332***	0.332***	0.322***	0.340***	0.320***
				(0.045)	(0.045)	(0.055)	(0.056)	(0.045)	(0.057)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 iterations) in parentheses. OLS regressions for equation (12). Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted when sample sizes are too small.*

while column (3) contains location fixed effects and examines how agents sort together within locations. In most cases the coefficients are positive and significant, suggesting that even within industries or within locations, the best workers team up with the best workers in other layers of firms. Column (4) reports regression results with both industry and location fixed effects. The findings indicate that there is positive assortative matching. For example, in organizations with four layers, a one hundred percent increase in the average ability of workers in layer one is associated with 0.322 average increase in the average ability of workers in layer two. In standardized units, this implies that a one standard deviation increase in the average ability of agents in layer one corresponds to a 0.301 standard deviation increase in the average ability of agents in layer two.

Column (5) in Table 4 reports results of regressions with firm observables as controls. Almost all of the coefficients are positive and half are significant. Finally column (6) reports results with the full set of controls, industry, location and firm observables. The coefficients remain positive and significant. For example, in organizations with three layers, a one hundred percent increase in the average ability of workers in layer one is associated with 0.237 average increase in the average ability of workers in layer two, and with a 0.292 average increase in the ability of workers in layer three. To obtain a sense of the strength of this relationship, a one standard deviation increase in the average ability of agents in layer two, corresponds to a 0.224 and 0.259 standard deviation increase in the average ability of agents in layers two and three, respectively.

One observation from Table 4 is that the magnitudes of α_1 are small. A small magnitude, however, is not necessarily inconsistent with the theory, since the assignment of agents into teams depends on the parameters of the model, and in particular on the distribution of abilities in the economy.[33] Furthermore, nearly all of the coefficients reported in Table 4 have a positive sign. There are two notable exceptions. In row six the coefficients between the ability of workers in layers four and two in four-layer firms are negative but not significant. In both cases, however, the reported coefficients are imprecise. Because the sample sizes are small relative to the number of control variables there is not much independent variation in the data, which may also account for the reported negative coefficients. In both cases results are inconclusive. In addition, there are several rows in Table 4 where although the coefficients are positive, they are never significant. For example, in organizations with four layers, there appears to be no relationship between the average ability of agents in layers three and one. The same results hold for two-layer organizations. This suggests that there is no sorting between agents in these layers, however the fact that the coefficients are always positive indicates that there is a relationship in the data, albeit not strong.[34]

To summarize the results, out of the possible 50 estimated coefficients, 28 are positive and significant at the five percent level, 3 are positive and significant at the ten percent level, 17 are positive but not significant, and 2 are negative and not significant. Therefore, apart for two-layer firms, these results provide evidence that there is positive assortative matching between workers in different layers of firms.

6.3 Examining the Mechanism

Until now, I have found evidence that higher ability agents sort into the higher layers of firms and that there is positive assortative matching between the workers in the different layers of firms. I now proceed to examine whether the model's mechanism is behind the sorting pattern observed

[33]The small magnitudes for α_1 are not problematic. If one were to assume a continuum of agents, as in Antras et al. (2006), then the mass of managers will be smaller than the mass of production workers. In this case, the matching function would have a slope that is less than 1.

[34]For two-layer firms, this is consistent with the findings in table 3, which report that there is little evidence of sorting between agents and layers in firms.

in the data and test whether agents' span of control increases with their subordinates' ability.[35]

Let HR_j^l denote the total number of hours worked by all employees in layer l at firm j. I define the span of control of workers in layer l as:

$$span_j^l = \frac{HR_j^1}{HR_j^l}. \tag{13}$$

In words, my measure of the span of control of workers in layer l is the ratio of the total number of hours in layer 1, to the number of hours in layer l.[36] The argument is that all workers in layer l supervise N_j^l individuals in layer 1, and these individuals spend a total of HR_j^1 hours at the firm. Dividing by the total number of hours worked by employees in layer l, HR_j^l, one obtains the number of hours a worker in layer l is expected to devote to supervising individuals in layer 1. This definition of span of control is directly related to firms' time constraint and is invariant to the number of hours in the highest layer of the organization.[37]

In the model, managers benefit from working with abler subordinates because they require less supervision, which allows managers to supervise more workers. In equilibrium, abler managers form teams with abler subordinates and their span of control is increasing with their subordinates' ability. To examine whether the model's mechanism holds in the data I estimate the following equation:

$$\ln span_j^l = \gamma_0 + \gamma_1 ability_j^{l-1} + X_j\beta + u_j, \tag{14}$$

where $ability_j^{l-1}$ is the estimated weighted average ability of all workers in firm j who are in layer $l-1$, and $span_j^l$ is defined above. The controls X_j are firm age, whether the firm was present in 1976, and indicator variables for industry and location. I estimate equation (14) for firms with the same number of layers and for different values of l, separately. In equation (14) the interest is in how the span of control of agents in layer l varies with the ability of agents in the layer below. If there is positive assortative matching between workers in the different layers of firms, and abler subordinates render their superiors more productive by allowing them to supervise more workers, then γ_1 should be positive and significant.

Table 5 reports regression results. The table has a similar structure to Table 4. Each entry in the table reports an estimated value of the coefficient γ_1. The first column reports the total number of layers in firms. The second column reports the layer for which weighted average ability is the right-hand-side variable in equation (14), the third column reports the sample size

[35]In the appendix I also examine whether agents' span of control increases with their ability. Given that both approaches examine the same question, and the findings are more conclusive when I examine how agents' span of control varies with their subordinates' ability, I relegate the latter analysis to the appendix.

[36]Since I cannot observe reporting relationships within organizations, this is the only measure available. I obtain HR_j^l from the exhaustive cross-section of the DADS.

[37]In the Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) all workers have one unit of time available. Also, the number of workers in the top layer of a firm is normalized to one.

Table 5: Testing Mechanism - Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	(1)	(2)	(3)	(4)	(5)	(6)
TWO	1	2863	-0.070	-0.098^{*}	-0.031	-0.068	-0.055	-0.061
			(0.055)	(0.054)	(0.058)	(0.056)	(0.055)	(0.056)
THREE	1	6430	-0.077^{**}	-0.035	-0.060^{*}	-0.032	-0.070^{*}	-0.028
			(0.035)	(0.034)	(0.036)	(0.035)	(0.036)	(0.035)
THREE	2	2413	-0.271^{***}	-0.173^{**}	-0.206^{**}	-0.158^{*}	-0.256^{***}	-0.149^{*}
			(0.080)	(0.075)	(0.086)	(0.084)	(0.079)	(0.084)
FOUR	1	4494	-0.143^{***}	-0.105^{**}	-0.115^{***}	-0.088^{**}	-0.140^{***}	-0.091^{**}
			(0.044)	(0.042)	(0.044)	(0.042)	(0.045)	(0.043)
FOUR	2	1918	-0.161^{**}	-0.089	-0.084	-0.067	-0.158^{**}	-0.067
			(0.081)	(0.073)	(0.085)	(0.079)	(0.080)	(0.078)
FOUR	3	1042	0.198	0.273^{**}	0.249^{*}	0.316^{**}	0.218^{*}	0.334^{**}
			(0.133)	(0.124)	(0.147)	(0.143)	(0.131)	(0.143)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regressions for equation (14). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (14). The left-hand side variable is the span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

of the regressions, and the fourth to ninth columns report estimated values of γ_1.[38] In Table 5 all the standard errors are bootstrapped and based on 500 replications.

Column (1) contains no controls and examines how span of control varies with the ability of agents in the layer below across industries and locations. In these regressions the majority of the estimated values of γ_1 are negative, indicating that the model's mechanism does not hold in the data and span of control is decreasing with workers' ability. For example in three-layer firms a one unit increase in the average ability of workers in layer two is on average associated with a 27.1 percent decrease in the span of control of workers in layer three.

Within industries and locations, there is evidence that abler agents limit their managers' span of control but there is also evidence in favor of the model's mechanism. Column (2) contains industry fixed effects and examines how agents' span of control varies with their subordinates' ability within industries, while column (3) examines the relationship within locations. In both models the majority of regressions indicate abler workers limit their managers' span of control.

[38] Note that the reported sample sizes in Table 5 vary across regressions. For example, in regressions with four-layer firms the sample size can be as small 1,042 observations or as large as 4,494 observations. The sample size also increases in regressions examining the mechanism in the lower layers of firms. As explained previously, this is not surprising given nature of the dataset and, because firms are hierarchies in the dataset there are less employees in the higher layers of firms. Also the reported sample sizes are different from Table 4, since to estimate equation (14) only the ability of one employee in a layer has to be recorded in the dataset.

The only exception is in the last row of the table, where in four-layer firms there is evidence that abler workers allow their managers to supervise larger teams. For example, within industries in four-layer firms a unit increase in the average ability of workers in layer three is associated with a 27.3 percent increase in the span of control of workers in layer four. Column (4) reports results with both location and industry controls. The results remain the same: apart for layer four in four-layer firms the evidence indicates that abler agents supervise smaller teams.

Column (5) reports results with firm observables as controls. In column (5) four coefficients are negative and significant at the ten percent level while one is positive and significant. Finally column (6) reports results with the full set of controls, industry, location and firm observables. With the exception of the last row, there is evidence that abler agents limit the number of workers managers' supervise. For example in three-layer firms a one unit increase in the average ability of workers in layer two is associated with a 14.9 percent decrease in the span of control of workers in layer three. In four-layer firms the relationship however remains positive and significant, and indicates that a one unit increase in the average ability of workers in layer three is associated with a 33.4 percent increase in the span of control of workers in layer four. In light of these results, only in layer four in organizations with four layers, is there evidence to suggest that the model's mechanism is present in the data.

To summarize the results reported in Table 5, out of the 36 estimated coefficients, 13 are negative and significant at the five percent level, 5 are negative and significant at the ten percent level, 12 are negative but not significant, 5 are positive and significant at the ten percent level, and 1 is positive but not significant. Hence overall the evidence is mixed. There is limited evidence in favor of the mechanism described by the model, and evidence to suggest that the opposite is taking place: abler managers form teams with abler production workers, however, these production workers take up more of the managers' time, which limits the amount of agents managers can supervise.

Additionally, as long as all agents perform the same tasks in firms, the size of each layer is determined by the same constraint and better workers form teams with better workers in the other layers of firms, it is incompatible for higher ability managers to supervise less workers and for higher ability workers to be employed in larger teams.[39] These results therefore indicate that it may be the case that different mechanisms are behind the sorting pattern observed in the data, and that firms are organizing production in different ways.[40] In some firms, abler managers supervise a greater number of abler workers, while in other firms abler managers supervise a smaller number of abler workers. I next section I explore whether there is any heterogeneity in the way firms organize production.

In contrast, it may also be the case that agents in the different layers of firms do not perform the same tasks or the size of each layer is determined by a different constraint, and so it may be

[39]For example, in Garicano (2000) all agents solve problems to produce output and so the size of each layer is determined by the same time constraint.

[40]In other words, the regressions thus far have only been estimating an average correlation across firms.

the case that the relationship between managers' span of control and their subordinates' ability varies within the different layers of firms. For some layers better subordinates may allow their managers to supervise more workers, while in other layers better subordinates may limit the number of workers their managers can supervise. For example because of the different tasks agents perform, in a four-layer firm it may be the case that in layers one and two managers' span of control decreases with their subordinates' ability, while in layers three and four managers' span of control increases with their subordinates' ability. To date, there has been no theoretical model that can justify such an organizational structure and so there are not any empirical predictions that I can examine in the data.

6.4 Additional Results

Together the results in the previous section suggest that, although there is evidence of positive assortative matching between workers in the different layers of firms, not all firms are organizing their production in a uniform way. To investigate whether this is the case, in this section I examine the data more closely. I proceed in three steps. First, I examine whether the relationships hold for mono-establishment firms, that is firms that consist of only one plant. Second, I examine whether the findings in the previous section are the same across industries with different degrees of technological intensity. And third, I examine whether the findings are the same across industries with different degrees of product differentiation. In this section I only report results for regressions that examine whether the model's mechanism is behind the sorting pattern observed in the data with the full set of controls, industry, location and firm observables. To provide a complete analysis of the data, I report regression results that examine whether managers' span of control increases with their subordinates' ability as well as results that examine whether managers' span of control increases with their own ability.[41] Results for regressions that examine how agents sort into layers and how agents sort together into teams are in the appendix.

First, it may be the case that reporting relationships are only specific to a physical location. In particular if a firm is operating multiple plants, their organization may be different than what is suggested by the theory. To account for this, in Table 6 I report regression results for mono-establishment organizations.[42] Column (1) reports regression results that examine how the span of control of workers in layer l varies with the average ability of workers in layer $l-1$, while column (2) reports results that examine how the span of control of workers in layer l varies with their average ability. In Table 6 the results are similar to those reported in the previous section. Apart for the span of control of agents in layer four, the estimated values of γ_1 are negative. Therefore, the conclusion remains the same. There is limited evidence in favor of the mechanism described by the model, and evidence to suggest that the agents' span of control is decreasing with ability.

[41] Given that there is positive assortative matching, both types of regressions should yield similar results.

[42] This removes 924 firms from the dataset.

Table 6: Testing Mechanism - Mono-Establishment Firms

Total Number of Layers	layer l-1	Sample Size	(1)	Total Number of Layers	layer l	Sample Size	(2)
TWO	1	2723	-0.072 (0.058)	TWO	2	591	-0.135 (0.192)
THREE	1	6140	-0.033 (0.038)	THREE	2	2309	-0.046 (0.078)
THREE	2	2309	-0.161^{*} (0.085)	THREE	3	1120	-0.075 (0.112)
FOUR	1	4203	-0.100^{**} (0.048)	FOUR	2	1783	-0.061 (0.066)
FOUR	2	1783	-0.058 (0.083)	FOUR	3	952	-0.102 (0.131)
FOUR	3	952	0.270^{*} (0.154)	FOUR	4	82	

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regressions for equations (15) and (14) . Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equations with the full set of controls. The left-hand side variable is the span of control of agents in layer l. In column (1) the right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - 1$. In column (2) the right-hand side variable is the estimated weighted average ability of workers in layer l. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted when sample sizes are too small.*

Second, it may be the case firms using different technologies organize their production differently. To account for this, I use the OECD (2003) classification of technological intensity to group industries into four categories of technological intensity, low, medium-low, medium-high and high, and re-estimate equation (14) with the right-hand side variable interacted with the technological intensity of the industry.[43] Table 7 report results. From Table 7 a weak pattern emerges. When it is significantly different from zero, in high technology industries γ_1 is negative, while in medium-low technology industries the opposite is the case. For example across organizations with three layers, in medium-low technology firms a unit a one unit increase in the average ability of workers in layer two is associated with a 25.5 percent increase in the span of control of workers in layer three, while in high technology firms it is associated with a 49.6 percent decrease. Furthermore, in Table 8 I conduct the same analysis but replace the right-hand side variable with the average ability of workers in layer l. A similar pattern is also present in Table 8. For example, across three-layer organizations, in medium-low technology firms a one unit increase in the average ability of workers in layer two is on average associated with a 29.8 percent increase in their span of control, while in high technology firms it is associated with a 58.2 percent decrease. Therefore, there is some weak evidence suggesting that the mechanism described by the model holds in firms operating in medium to low technology industries, while in high technology industries the evidence suggests that abler agents supervise less workers.

[43]See Appendix A for further details.

Table 7: Testing Mechanism - Technology Intensity: Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	Low	Medium-Low	Medium-High	High
TWO	1	2674	-0.074	-0.073	-0.066	0.434
			(0.085)	(0.097)	(0.152)	(0.303)
THREE	1	5868	-0.045	0.013	-0.086	-0.028
			(0.058)	(0.057)	(0.092)	(0.200)
THREE	2	2208	-0.200	0.255^{*}	-0.115	-0.496^{*}
			(0.135)	(0.149)	(0.180)	(0.298)
FOUR	1	4109	-0.112	-0.060	-0.137	-0.185
			(0.069)	(0.074)	(0.096)	(0.190)
FOUR	2	1764	-0.058	0.027	-0.090	-0.505^{*}
			(0.155)	(0.158)	(0.161)	(0.258)
FOUR	3	953	0.398	0.095	0.181	-0.202
			(0.270)	(0.286)	(0.313)	(0.369)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regressions for equation (14). Each row displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 interacted with industries degree of technological intensity, for regressions with the full set of controls. The left-hand side variable is the span of control of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

Third, firms' organization may also depend on the type of product they are producing. It may be the case that firms producing differentiated products are required to organize differently than firms producing homogeneous products, either because the production process is different across goods, or because the level of product market competition is different across markets. To account for this, I use the Rauch (1999) classification of goods to group industries into two categories of product differentiation, low and high, and re-estimate equations with the right-hand side variable interacted with the degree of production differentiation of the industry.[44] Table 9 report results and in Table 9 a very weak pattern emerges. In the few cases where it is significantly different from zero, γ_1 is negative in highly differentiated industries, while in industries with a low degree of product differentiation γ_1 is positive. For example across four-layer organizations, in industries with a low degree of product differentiation a one unit increase in the average ability of workers in layer three is on average associated with a 58.8 percent increase in the span of control of workers in layer four, while in industries with a high degree of product differentiation a unit increase in the average ability of workers in layer one is associated with a 12.3 decrease in the span of control of workers in layer two. Therefore, there is some weak evidence suggesting that the mechanism described by the model holds in firms operating in non-differentiated industries, while in highly differentiated industries the evidence suggests that abler agents supervise less workers.

[44]See Appendix A for further details.

Table 8: Testing Mechanism - Technology Intensity: Managers' Ability

Total Number of Layers	layer l	Sample Size	Low	Medium-Low	Medium-High	High
TWO	2	571	-0.154	-0.091	0.055	-0.142
			(0.303)	(0.374)	(0.456)	(0.420)
THREE	2	2208	0.034	0.298^{**}	-0.105	-0.582^{**}
			(0.126)	(0.143)	(0.178)	(0.251)
THREE	3	1090	0.190	-0.227	-0.114	-0.586^{**}
			(0.184)	(0.202)	(0.289)	(0.289)
FOUR	2	1764	-0.042	0.067	-0.149	-0.159
			(0.120)	(0.131)	(0.136)	(0.229)
FOUR	3	953	0.089	-0.084	-0.349	-0.070
			(0.232)	(0.183)	(0.255)	(0.335)
FOUR	4	76				

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regressions for equation (15). Each row displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 interacted with industries degree of technological intensity, for regressions with the full set of controls. The left-hand side variable is the span of control of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in layer l. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted when sample sizes are too small.*

7 Robustness Checks

7.1 Inconsistent Estimates

There are three threats to my estimates of workers' ability. All stem from my estimation of worker fixed effects. First, because the worker fixed effects are incidental parameters from regression (9), consistent estimates for them can only be obtained as the number of years an individual is observed in the panel grows large. Since for the years 1993 to 2004, the average worker is observed for 6 years, not all of the estimates of the time-invariant component, θ_i, identify a consistent measure of a worker's ability, which in turn may lead to measurement error in my estimate of the average ability of workers in a layer of a firm. Although the panel is short, to get a sense of how important is this issue, I conduct my analysis on workers that I observe for at least 10 periods. Further, because I have established that the mechanism driving the sorting pattern is heterogeneous across firms, in this section I only report results for skill stratification and positive assortative matching. Tables A18 and A19 in the appendix present the regression results for this restricted sample.

Table A18 reports tests for skill stratification. As in the previous table, higher ability agents occupy the upper layers of organizations. In addition in column (5), even within two-layer firms this relationship is now significant and indicates that an individual with a one hundred percent increase in his ability will on average reside 0.254 layer higher.

Table 9: Testing Mechanism - Product Differentiation

Total Number of Layers	layer l-1	Sample Size	Low	High	Total Number of Layers	layer l	Sample Size	Low	High
TWO	1	2858	0.001	-0.061	TWO	2	599	-0.222	-0.148
			(0.099)	(0.066)				(0.598)	(0.206)
THREE	1	6409	-0.043	-0.018	THREE	2	2399	0.158	-0.042
			(0.080)	(0.044)				(0.141)	(0.089)
THREE	2	2399	-0.198	-0.060	THREE	3	1178	-0.022	-0.103
			(0.148)	(0.099)				(0.184)	(0.123)
FOUR	1	4490	-0.029	-0.123^{**}	FOUR	2	1917	-0.113	-0.018
			(0.095)	(0.080)				(0.126)	(0.072)
FOUR	2	1917	-0.198	-0.0009	FOUR	3	1041	-0.079	-0.102
			(0.142)	(0.091)				(0.277)	(0.132)
FOUR	3	1041	0.588^{*}	0.224	FOUR	4	82		
			(0.335)	(0.162)					

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equations with the full set of controls. The left-hand side variable is the span of control of agents in layer l. In column (1) the right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$, interacted with industries degree of product differentiation. In column (2) the right-hand side variable is the estimated weighted average ability of workers in layer l. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted when sample sizes are too small.*

Table A19 reports regression results that test for positive assortative matching. First, note the results are similar to Table A19 and so there is strong evidence in favor of positive assortative matching between workers in the different layers of firms. Second, in contrast to the results reported in Table 4, there is now weak evidence indicating a positive correlation between workers' ability in layers two and one in two-layer firms and layers three and one in four-layer firms. In column (2), for example, a unit increase in the weighted average ability of workers in layer one corresponds to a 0.137 increase in the average ability of workers in layer three. Hence, when I remove inconsistent estimates of worker's ability the general conclusions remain the same: the evidence indicates that higher ability agents occupy the upper layers of firms and that there is positive assortative matching between workers in the different layers of firms.

7.2 Estimation Error

A positive correlation between the individual fixed effects may be the outcome of using standard econometric techniques. As discussed in Abowd et al. (2004) and Andrews et al. (2008), in equation (9) there is a negative correlation between the worker and firm-layer effects caused from standard estimation error. When the firm-layer fixed effects in equation (9) are on average underestimated, the individual fixed effects will be overestimated, and when the firm-layer fixed effects are on average overestimated, the individual fixed effects will be underestimated. Because in the panel workers transition between layers within firms, this implies that my regressions may suffer from non-classical measurement error, biasing results.

To resolve these issues I conduct my analysis only on workers who have moved to a new employer in the year 2009. For this sample of workers, any errors caused by miss-estimated firm-layer fixed effects will be uncorrelated with one another. In these regressions, the coefficients will only suffer from attenuation bias, however the sign of the estimated coefficients will more properly reflect the relationships of interest. Because the sample sizes are small in these regressions, unlike in the previous sections, Tables A20 and A21 do not report results from all models.

For the year 2009, Table A20 reports how workers sort into layers. In all regressions the coefficients of interest are positive, and indicate the higher ability workers occupy the upper layers of firms. The findings are similar to Table 3. In two-layer firms, except when I examine how workers sort into layers within firms, the coefficients are significant at the one percent level. In addition in three and four-layer firms all the reported coefficients are significant at the one percent level. For example, the findings indicate that within a three-layer firm an individual with a unit increase in his ability will on average reside 0.411 layers higher. Therefore the conclusion remains the same, higher ability agents sort into higher layers in firms.

Table A21 reports results of tests examining whether better workers are employed with better workers in the other layers of firms. Because there are not many workers employed in the same firm, Table A21 only reports results for three-layer and four-layer organizations. In firms with three layers there is evidence of positive assortative matching. The majority of the reported coefficients have a positive sign, and five coefficients are significant at the five percent level. For example, across industries and locations a one unit increase in the average ability of agents in layer two corresponds to a 0.368 increase in the average ability of workers in layer three. Because the sample sizes are relatively small, as additional controls are included these relationships remains positive but are no longer significant. In organizations with four layers, there is evidence to suggest that there is a positive correlation between workers' ability in the different layers of firms. The majority of the reported coefficients have a positive sign, and six are significant at the five percent level. For example, within industries and locations, a one unit increase in the average ability of agents in layer one corresponds to a 0.291 average increase in the average ability of workers in layer two. Overall the findings are consistent with previous results and the conclusion remains the same: there is evidence indicating that better workers are employed with better workers in the other layers of firms.

7.3 Biased Estimates

Third, the estimated worker fixed effects may be biased. If workers in a given layer render their subordinates more productive, and if their subordinates make them more productive, then this should be reflected in wages. If this is the case, then for each individual the worker fixed effects in equation (9) are not only identifying the ability of a worker, but also the average impact his co-workers have on his ability. Hence, this would imply that there is non-classical measurement error in my estimates of workers' ability, which may explain why there is a positive correlation between the ability of workers in the different layers of firms.[45]

Biased estimates are a concern for tests that examine whether abler managers occupy the upper layers of firms and tests of positive assortative matching between workers in the different layers of firms.[46] This issue was implicitly addressed in several parts of the paper. First, any time-invariant complementarities between workers in the layers of firms are accounted for by the firm-layer fixed effects. Second, the results the Section 5 also indicate even though there is a another component to wages, equation (9) is a modest approximation to wages.[47] Finally, because it focuses on workers who move to a new firm, the analysis in Section 7.2 partially addresses this concern. In this case, the worker fixed effects are uncorrelated with one another. The results in Section 7.2 are consistent with the analysis in the paper: there is evidence that abler managers occupy higher layers in firms and evidence in favor of positive assortative matching between workers in the different layers of firms.

8 Conclusion

How workers sort together with other workers into layers and firms is crucial for understanding the organization of firms. Without knowledge of the precise nature of the interactions between workers in the different layers of firms it is difficult to comprehend how firms organize production. Better knowledge of how workers sort into layers and firms is also important for understanding earnings' inequality, how firms respond to changes in their market environment, and how labor is allocated in the aggregate economy. Therefore identifying the mechanism that is causing this sorting pattern is essential for understanding the nature of firms.

This paper directly examined how workers sort together in firms. My empirical strategy relies on the idea that firms can be thought of as hierarchical teams, composed of layers that perform different tasks. Using French administrative data, I classify employees into organizational layers

[45]In addition, if his co-workers have on a workers' earning, this would imply that all my estimates in regressions (12), (15) and (14) are an upper bound to the actual relationship of interest.

[46]Moreover if the model is an accurate description of the real world, since there is positive assortative matching between workers in the different layers of firms, for a given layer, a peer effect should be positively correlated with the ability of a worker. Therefore, according to the model the bias should be increasing in the ability of an individual. For regressions (12), (15) and (14), this would further bias the coefficient of interests, α_1 and γ_1 in favor of finding a positive result. Therefore, one interpretation of my results is that they present an upper bound to the relationships of interest.

[47]Returning to the case of a worker employed in layer 1 in Firm 1 in period t who moves to layer 1 in Firm 2 in period $t+1$. Let $I(j,t)$ denote the set of workers in firm j in period t. Ignoring the returns to observables, and assuming a peer effect, $\psi(i, I(j,t))$, his expected change in wages is equal to: $E[w_{it+1} - w_{it} | J(i,l,t+1) = \{1,2\}, J(i,l,t) = \{1,1\}] = \psi(1, I(2,t+1)) - \psi(1, I(1,t)) + \phi_{1,2} - \phi_{1,1} + E[\epsilon_{it+1} - \epsilon_{it} | J(i,l,t+1) = \{1,2\}, J(i,l,t) = \{1,1\}]$. Similarly, the expected change in wages of a worker moving in the other direction is equal to: $E[w_{it+1} - w_{it} | J(i,l,t+1) = \{1,1\}, J(i,l,t) = \{1,2\}] = \psi(2, I(1,t+1)) - \psi(2, I(2,t)) + \phi_{1,1} - \phi_{1,2} + E[\epsilon_{it+1} - \epsilon_{it} | J(i,l,t+1) = \{1,1\}, J(i,l,t) = \{1,2\}]$. Therefore if the error term is strictly exogenous both expressions are simply equal to $\psi(1, I(2,t+1)) - \psi(1, I(1,t)) + \phi_{1,2} - \phi_{1,1}$ and $\psi(2, I(1,t+1)) - \psi(2, I(2,t)) + \phi_{1,1} - \phi_{1,2}$.

as in Caliendo et al. (2014). I conclude that, within firms, higher ability workers are employed in the higher layers of firms, and across firms, there is positive assortative matching between workers in the different layers of firms. However I find only weak evidence for the mechanism, as suggested by Garicano and Rossi-Hansberg (2006), to be causing this sorting pattern: higher ability workers allow their managers to increase their span of control and employ larger teams. Finally, I also find evidence that the opposite is taking place: higher ability workers limit managers' span of control.

An important question remains to be answered. The findings presented in this study indicate that the nature of interactions between workers in the different layers of firms are many and diverse, and firms are not organizing their production in a uniform way. Although there is some evidence that higher ability workers allow their managers to increase their span of control and employ larger teams, there is also evidence that the opposite is taking place. Some of the organizational differences of firms can be explained by their degree of technological intensity and the products they produce. An important question however still remains to be answered: If better workers sort into firms with other better workers, what is causing this sorting pattern?

References

Abowd, John M., Creecy Robert H., and Francis Kramarz. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." *Working Paper* .

Abowd, John M., Francis Kramarz, Paul Lengerman, and Sebastien Perez-Duarte. "Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States." *Working Paper* .

Abowd, John M., Francis Kramarz, and David N. Margolis. "High Wage Workers and High Wage Firms." *Econometrica* 67, 3: (1999) 251–333.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Upward Richard. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Statistical Artefact?" *Journal of the Royal Statistical Society* 171, 3: (2008) 673–697.

Antras, Paul, Luis Garicano, and Esteban Rossi-Hansberg. "Offshoring in a Knowledge Economy." *Quarterly Journal of Economics* 121, 1: (2006) 31–77.

———. *Organizing Offshoring: Middle Managers and Communication Costs*. Cambridge, US: Harvard University Press, 2008, e. helpman, d. marin, t. vierdier eds. edition.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. "Anatomy of French Production Hierarchies." *Journal of Political Economy* Forthcoming.

Card, David, Jorg Heining, and Patrick Kline. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128, 3: (2013) 967–1015.

Coase, Ronald. "The Nature of the Firm." *Economica* 16, 4: (1937) 386–405.

Dasgupta, Kunal. "Learning and Knowledge Diffusion in a Global Economy." *Journal of International Economics* 87, 2: (2012) 323–336.

Eeckhout, Jan, and Philipp Kircher. "Identifying Sorting - In Theory." *Review of Economic Studies* 78, 1: (2011) 872–906.

———. "Assortative Matching with Large Firms: Span of Control over More versus Better Workers." *Working Paper* .

Garicano, Luis. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economics* 108, 5: (2000) 874–904.

Garicano, Luis, and Thomas N. Hubbard. "Hierarchical Sorting and Learning Costs: Theory and Evidence from the Law." *Journal of Economic Behavior and Organization* 58, 2: (2005) 349–369.

———. "The Returns to Knowledge Hierarchies." *Working Paper* .

Garicano, Luis, and Esteban Rossi-Hansberg. "Inequality and the Organization of Knowledge." *American Economic Review Papers and Proceedings* 50, 1: (2004) 197–202.

———. "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121, 4: (2006) 1383–1435.

Guimaraes, Paulo, and Pedro Portugal. "A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects." *The Stata Journal* 10, 4: (2010) 628–649.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. "Skill Dispersion and Firm Productivity: An Analysis with Matched-Employer Employee Data." *Journal of Labor Economics* 26, 2: (2008) 247–285.

Lucas, Robert E. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9, 2: (1977) 508–523.

Martins, Pedro S. "Dispersion in Wage Premiums and Firm Performance." *Economic Letters* 101, 2: (2008) 63–65.

de Melo, Rafael Lopes. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Working Paper* .

OECD. *OECD Science, Technology and Industry Scoreboard*. OECD, 2003.

Rauch, James E. "Networks versus Markets in International Trade." *Journal of International Economics* 48, 1: (1999) 7–35.

Rosen, Sherwin. "Authority, Control, and the Distribution of Earnings." *The Bell Journal of Economics* 13, 2: (1982) 311–323.

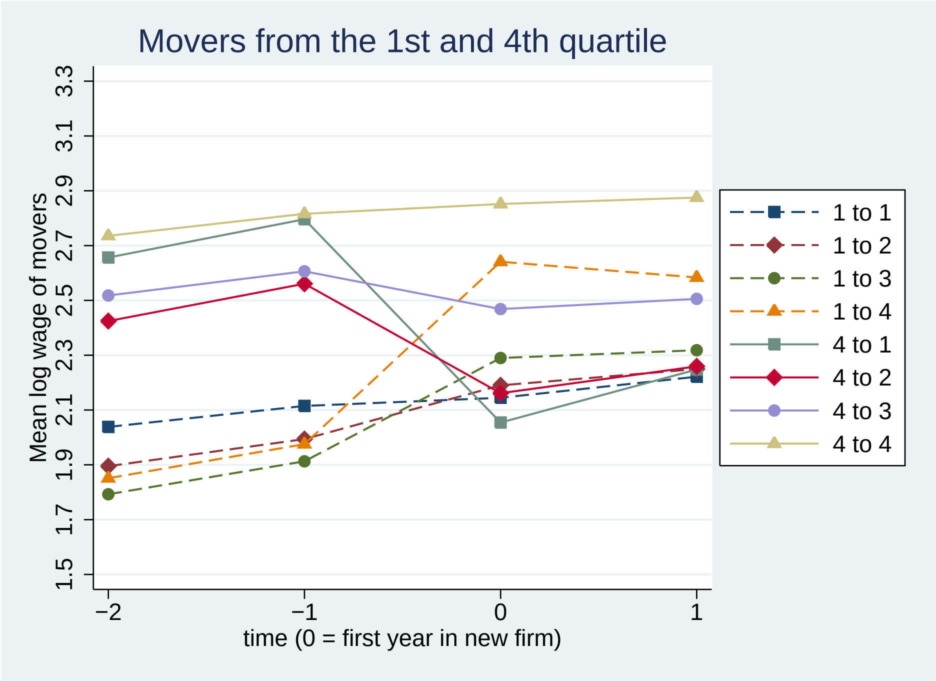
Appendix A: Data Appendix

Description of Results

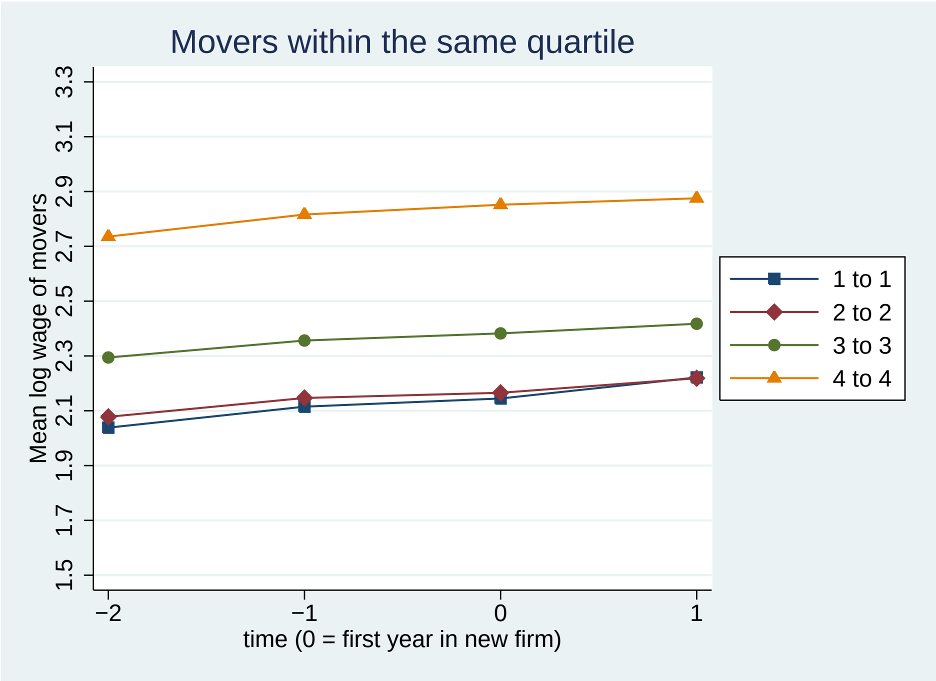
Table A1: Mean Log Wages by Transitions and Years

Origin-Destination Quartile	Number of Observations	Two Years Before	One Year Before	One Year After	Two Years After	Change from Two Years Before and After
1 to 1	25,775	2.03	2.11	2.14	2.22	0.19
1 to 2	15,654	1.89	1.99	2.19	2.24	0.35
1 to 3	11,759	1.79	1.91	2.28	2.31	0.52
1 to 4	3,410	1.85	1.97	2.64	2.58	0.73
2 to 1	16,427	2.09	2.20	2.03	2.12	0.03
2 to 2	51,732	2.07	2.14	2.16	2.21	0.14
2 to 3	44,670	2.07	2.16	2.28	2.32	0.25
2 to 4	10,489	2.08	2.19	2.55	2.55	0.47
3 to 1	11,468	2.25	2.37	1.98	2.11	−0.14
3 to 2	40,717	2.17	2.25	2.14	2.21	0.04
3 to 3	109,545	2.29	2.35	2.38	2.41	0.12
3 to 4	42,056	2.42	2.50	2.66	2.69	0.27
4 to 1	3,445	2.65	2.79	2.05	2.24	−0.41
4 to 2	8,550	2.42	2.56	2.16	2.25	−0.17
4 to 3	30,478	2.51	2.60	2.46	2.50	−0.01
4 to 4	72,529	2.73	2.81	2.85	2.87	0.14

Notes: Descriptive statistics of job transitions from the estimation of equation (9).



(a) Movers from 1st and 4th Quartile



(b) Movers within the Same Quartile

Figure 2: Wages of Movers

The Panel Dataset of the DADS

To estimate worker and firm fixed effects, I use the years 1993 to 2004 from the panel dataset of the DADS. Initially, the dataset contains 24,882,933 total observations, 5,469,362 workers and 1,614,337 firms. I remove from the dataset any workers or firms that cannot be properly identified or that have missing values. For reasons of computational tractability, I restrict the sample to all workers who are born in an even numbered year, are between the ages 18 and 65 and work in continental France. I also eliminate from the sample all individuals I observe only once in the panel and who are not full-time workers. In a given year, an individual may hold multiple jobs. In case of multiple jobs, for a given year I keep the worker's employment with the highest salary. Finally, I also eliminate all firms in the agricultural and fishing industries. From this sample of workers and firms, to obtain an exact estimate of worker and firm-layer fixed effects I find the largest connected group. The largest connected group contains 4,999,728 observations, 753,092 workers, 399,676 firms and 569,198 firm-layer pairs.

For the years 1993 to 2004 Table A2 presents distribution of the number of years workers are observed in the panel dataset of the DADS.

Table A2: Distribution of the number of years workers are observed in the panel

variable	mean	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Years	6.63	2	2	2	2	4	6	10	11	11	12	12

For the years 1993 to 2004 Table A3 presents distribution of the number of years firms are observed in the panel dataset of the DADS as well as the distribution of the number of workers that are observed in a layer in a given year.

Table A3: Distribution of the number of years layers within firms are observed in the panel and the distribution of workers per layer in a firm

variable	mean	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Years	3.28	1	1	1	1	1	2	4	8	10	11	12
Workers	2.67	1	1	1	1	1	1	2	4	7	22	9,674

The Exhaustive Cross-Section of the DADS

The exhaustive cross-section of the DADS contains information on all workers who earn a positive wage in a french establishment. For a given year, the observations are at the worker-establishment level. To clean the data, I first remove any observations that do not have a positive amount of hours, days, or wages reported or that do not have an occupation. I also remove any observations in which the firm and individual identifiers are missing and firms employing workers with occupational codes different from 2 to 6.

For every establishment there is information on its location (341 employment areas), the industry it operates in and the parent firm. I aggregate the data to the level of the firm, and retain only firms where all establishments operate in the same industry. Within a firm, a worker can have multiple jobs if he is employed in two different establishments. In this case, I treat the worker as two separate observations.

To identify the layers in a firm, I use the first digit of the CS occupational codes which range from 2 to 6. Therefore in total I can identify up to four layers. Layer 1 corresponds to qualified and non-qualified administrative workers and blue-collar workers. It contains all workers with CS occupational codes 5 and 6, respectively. I group CS occupational codes 5 and 6 together because their distribution of ability are similar. Layer 2 is composed of supervisors and individuals with higher level of responsibility than ordinary workers, and contains all workers with an occupational code 4. Layer 3 is composed of senior directors and top management staff and contains all workers with an occupational code 3. Layer 4 corresponds to owners who receive a wage and CEOs. It contains all workers with a CS occupational code 2. Further I consider a firm as having a layer if at least one employee is present in that layer and classify firms by the total number of layers in their organization. For example, a firm where layers 1, 3 and 4 are present and a firm where layers 1, 2 and 4 are present are both considered three-layer firms.

Merged Datasets

For the year 2004 I merge the information from the panel and exhaustive cross-section datasets together. Unlike the exhaustive cross-section, since the panel data is based on a 5 percent sample of the French population, it contains information on a sample of all firms operating in mainland France. Approximately 1 percent of firms in the panel dataset are not matched. I keep only firms that operate in the non-service sectors, and remove any firms that operate in more than one industry and location. In total the matched dataset contains $23,916$ firms that operate in 17 industries, of which 2,160 are one-layer firms, 3,322 are two-layer firms, 7,860 are three-layer firms and 5,450 are four-layer firms. Table A4 reports number of firms and workers observed in the matched sample by industry.

Table A4: Number of Firms by Industry

Industry	Number of Firms	Number of Workers
Agricultural & Food	3,363	7,217
Apparel & Leather	657	1,533
Publishing & Printing	1,351	2,618
Pharmaceuticals & Perfumes	226	722
Domestic Appliances & Furniture	1,063	2,417
Automobiles	373	1,700
Ships, Aircraft, Railroad Equipment	235	748
Machinery	3,070	6,138
Electric and Electronic Equipment	852	1,651
Building Materials & Glass Products	733	1,394
Textiles	650	1,594
Wood & Paper	1,162	2,643
Chemicals, Rubber & Plastics	1,382	3,728
Basic Metals & Metal Products	3,014	6,752
Electric & Electronic Components	564	1,568
Combustibles	23	95
Water, Gas & Electricity	72	181

Tables A5 and A6 contains descriptive statistics of firms in the exhaustive cross-section dataset and the matched dataset for the year 2004, respectively. As is evident from the average and the median number of workers and the average and the median number of hours in a firm, the matched dataset contains larger firms.[48] For both measures of firm size, number of employees

[48]As explained further below, one reason for this result is that I obtain my measures of workers' ability from the

Table A5: Descriptive Statistics from the Exhaustive Cross-Section Dataset

Exhaustive Cross-Section:

Total		Average		Median		Standard Deviation	
Number of Layers	Number of Firms	Number of Employees	Number of Hours	Number of Employees	Number of Hours	Number of Employees	Number of Hours
1	160,904	3.45	3,382.28	2	2,028	4.73	4,980.97
2	74,676	8.36	9,437.35	5	6,084	10.80	11,971.35
3	52,949	23.90	31,701.27	11	13,342	53.59	80,918.29
4	14,434	59.92	82,872.88	33	47,106	96.12	119,415.10

Notes: Descriptive statistics from the exhaustive cross-section of the DADS for the year 2004.

Table A6: Descriptive Statistics from the Matched Dataset

Matched Sample:

Total		Average		Median		Standard Deviation	
Number of Layers	Number of Firms	Number of Employees	Number of Hours	Number of Employees	Number of Hours	Number of Employees	Number of Hours
1	2,160	11.00	11,844.12	8	8,233.50	23.14	25,319.03
2	3,322	19.88	25,989.84	14	18,173.50	22.11	28,176.84
3	7,860	67.14	101,683.20	37	53,741	115.92	183,040.00
4	5,450	83.61	128,090.50	53	81,350	99.54	155,456.00

Notes: Descriptive statistics from the the matched sample dataset for the year 2004.

and the number of hours, the standard deviation in the matched sample is also greater than in the population. The sample is therefore biased towards larger firms, and is not representative of the entire population of firms. To the extent that the theory applies to all firms the unrepresentativeness of the sample is not of concern.

Technological Intensity

To group industries by their degree of technological intensity, I use the first four digits of the French classification of industries, NAF Rev 1, which corresponds to ISIC Rev 3.1. I then apply the same grouping as in the OECD (2003) classification. The table below, reproduced from Annex 1 of the OECD (2003), presents the list of industries by their degree of technological intensity.

Table A7: Technological Intensity

	Low	Medium-Low	Medium-High	High
Industries	36 to 37	351	31	353
Industries	20 to 22	25	34	2453
Industries	15 to 16	23	24 exl. 2453	30
Industries	17 to 19	26	352 & 359	32
Industries		27 to 28	29	33

Notes: List of industries by their degree of technological intensity. Reproduced from the OECD (2003).

largest connected set of workers and firms.

Product Differentiation

To group industries by their degree of production differentiation, I first assign each product in the Rauch (1999) classification to at least one ISIC Rev 3.1 industry using ISIC Rev 2 as a crosswalk. Second, using the conservative definition of products, I assign a value of 1 to homogeneous goods, 2 to reference priced goods and 3 to differentiated products. Third, for each industry I then calculate the average degree of differentiation of goods mapped into an industry giving equal weight to each product. And fourth I define industries with a low degree of product differentiation as the industries where the average degree of differentiation is less than 2.5, and industries with a high degree of production differentiation as the industries where the average degree of differentiation is greater than or equal to 2.5. The table below presents the list of industries by their degree of product differentiation.

Table A8: Product Differentiation

	Low	High
Industries	1 to 2	10
Industries	5	14
Industries	11 to 13	18 to 20
Industries	15 to 17	22
Industries	21	25 to 26
Industries	23 to 24	28 to 36
Industries	27	72
Industries	37	
Industries	40	
Industries	52	

Notes: List of industries by their degree of product differentiation.

Appendix B: Test for Positive Assortative Matching

In this section I test whether there is positive assortative matching between workers and their co-workers in a firm. I adopt the empirical test proposed by de Melo (2013) who shows that even though wages are not monotone with respect to firm productivity, wages will be monotone with respect to workers' skills. Therefore, if better workers are sorting together into firms, the correlation between a worker's fixed effect and the average fixed effect of his co-workers should be positive.

To examine whether this prediction holds in the data, I conduct two tests. For the year 2004, I first correlate the worker fixed effect with the average fixed effect of his co-workers across all firms in the economy, across firms with the same organization, and across the different layers of firms. And second, for the 2009 I conduct the same tests on the sample of workers who have moved to a new firm.

Tables A9 and A10 presents the results. The first column of table A9 presents the sample size and the second column contains the correlation across all workers in the economy and across firms with the same number of layers. The third column and fourth columns contain the same estimates for the year 2009. In all cases, the correlations reported are positive. For example across all firms in the economy, the correlation between the worker fixed effect and the average fixed effect of his co-workers is 0.352. Table A10, has a similar structure as Table A10 and goes further and estimates the correlation within the different layers of firms. The conclusions are similar: all reported correlations are positive, and so there is evidence of positive assortative matching between workers in firms.

Table A9: Correlations Across Firms

Total Number of Layers	Year 2004 N	Year 2004 $corr(\theta_i; \overline{\theta}_{-i})$	Year 2009 N	Year 2009 $corr(\theta; \psi)$
All	$31,941$	0.352	$8,302$	0.221
ONE	481	0.374	144	0.153
TWO	$1,871$	0.382	395	0.168
THREE	$15,773$	0.362	$4,538$	0.227
FOUR	$13,816$	0.335	$3,225$	0.219

Notes: Correlations between the ability of workers and their co-workers across firms with the same number of layers.

Table A10: Correlations within Layers of Firms

Total Number of Layers	layer	Year 2004 N	Year 2004 $corr(\theta_i; \overline{\theta}_{-i})$	Year 2009 N	Year 2009 $corr(\theta_i; \overline{\theta}_{-i})$
ONE	1	481	0.374	144	0.153
TWO	1	$1,819$	0.361	332	0.182
TWO	2	52	0.639	23	0.403
THREE	1	$15,026$	0.355	$2,681$	0.191
THREE	2	648	0.310	535	0.348
THREE	3	99	0.539	342	0.198
FOUR	1	$13,127$	0.325	$1,829$	0.184
FOUR	2	550	0.300	435	0.314
FOUR	3	139	0.429	238	0.090
FOUR	4	–	–	–	–

Notes: Correlations between the ability of workers and their co-workers within the layers of firms.

Appendix C: Testing the Mechanism - Managers' Ability

In this section, I examine whether agents' span of control increases with ability. I estimate the following equation:

$$\ln span_j^l = \gamma_0 + \gamma_1 ability^l{}_j + X_j\beta + u_j, \tag{15}$$

where $ability^l{}_j$ is the estimated weighted average ability of all workers in firm j who are in layer l, and $span_j^l$ is defined above. The controls X_j are firm age, whether the firm was present in 1976, and indicator variables for industry and location. I estimate equation (15) for firms with the same number of layers and for different values of l, separately. In equation (15) the interest is in how the span of control of agents in layer l varies with their ability. If there is positive assortative matching between workers in the different layers of firms, and abler subordinates render their superiors more productive by allowing them to supervise more workers, then γ_1 should be positive and significant.

Table A11 reports regression results. The table has a similar structure to Table 4. Each entry in the table reports an estimated value of the coefficient γ_1. The first column reports the total number of layers in firms. The second column reports the layer for which weighted average ability is the right-hand-side variable in equation (15), the third column reports the sample size of the regressions, and the fourth to ninth columns report estimated values of γ_1.[49] In Table A11 all the standard errors are bootstrapped and based on 500 replications.

Column (1) contains no controls and examines how workers' ability varies with their span of control across industries and locations. The results indicate that workers' span of control is decreasing with their ability. For example in three-layer firms a one unit increase in the average ability of workers in layer two is on average associated with a 17.5 percent decrease in their span of control.

Within industries and locations, the relationship remains negative but not significant. Column (2) examines how workers' span of control varies with their ability within industries, while column (3) examines the relationship within locations. In both models the evidence suggests that there may be a negative relationship between agents' ability and their span of control, however this is not conclusive. Column (4) examines the relationship within industries and locations. The results remain the same: although the coefficients are negative, they are not significant.

Column (5) reports results with firm observables as controls. In column (5) two coefficients are negative and significant at the five percent level. In organizations with three layers, a one unit increase in the average ability of workers in layer three is associated with a 27.3 decrease in their span of control, while a one unit increase in the average ability of workers in layer two is associated with a 15.8 decrease in their span of control. Finally column (6) reports results with the full set of controls, industry, location and firm observables. The coefficients remain negative but not significant.

In Table A11 nearly all of the estimates are negative, however only a handful of them are significant. The only consistent exception is the regressions for agents in layer four in four-layer firms, reported in the last row of the table. In this case the coefficients alternate sign, however

[49] Note that the reported sample sizes in Table A11 vary across regressions. For example, in regressions with four-layer firms the sample size can be as small 80 observations or as large as 1,918 observations. The sample size also increases in regressions examining the mechanism in the lower layers of firms. As explained previously, this is not surprising given nature of the dataset and, because firms are hierarchies in the dataset there are less employees in the higher layers of firms. Also the reported sample sizes are different from Table 4, since to estimate equation (15) only the ability of one employee in a layer has to be recorded in the dataset.

Table A11: Testing Mechanism - Managers' Ability

Total Number of Layers	layer l	Sample Size	(1)	(2)	(3)	(4)	(5)	(6)
TWO	2	601	−0.243	−0.161	−0.213	−0.109	−0.249	−0.115
			(0.161)	(0.132)	(0.221)	(0.196)	(0.161)	(0.192)
THREE	2	2413	−0.175**	−0.083	−0.081	−0.037	−0.158**	−0.027
			(0.075)	(0.069)	(0.082)	(0.075)	(0.074)	(0.074)
THREE	3	1183	−0.232**	−0.117	−0.067	−0.041	−0.273**	−0.068
			(0.114)	(0.101)	(0.125)	(0.118)	(0.111)	(0.117)
FOUR	2	1918	−0.120*	−0.068	−0.062	−0.055	−0.117*	−0.056
			(0.064)	(0.060)	(0.068)	(0.062)	(0.063)	(0.062)
FOUR	3	1042	−0.139	0.001	−0.216	−0.109	−0.116	−0.092
			(0.126)	(0.107)	(0.133)	(0.118)	(0.123)	(0.118)
FOUR	4	82	0.090	−0.001	0.313		−0.074	
			(0.299)	(0.369)	(0.517)		(0.301)	
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regressions for equation (15). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (15). The left-hand side variable is the span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in layer l. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted when sample sizes are too small.*

they are never significant, and so they do not lead to a firm conclusion.

To summarize the results reported in Table A11, out of the 34 estimated coefficients, 4 are negative and significant at the five percent level, 2 are negative and significant at the ten percent level, 24 are negative but not significant, and 3 are positive but not significant. In light of these findings, the evidence suggests that although there is positive assortative matching between workers in the different layers of firms, the model's mechanism explaining this sorting pattern is not present in the data. Indeed, the findings suggest that the opposite may be taking place. Abler managers form teams with abler production workers, however these workers take up more of the managers' time.

Appendix D: Mono-Establishment Firms

Table A12: Mono-Establishment Firms: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	(1)	(2)	(3)	(4)	(5)
TWO	$4,233$	0.054***	0.043***	0.046***	0.040***	0.043
		(0.013)	(0.013)	(0.014)	(0.015)	(0.134)
THREE	$18,714$	0.370***	0.346***	0.350***	0.336***	0.411***
		(0.015)	(0.014)	(0.013)	(0.013)	(0.019)
FOUR	$14,591$	0.460***	0.438***	0.432***	0.423***	0.499***
		(0.017)	(0.017)	(0.016)	(0.016)	(0.021)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (10) for workers in mono-establishment firms. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France.*

Table A13: Mono-Establishment Firms: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	(1)	(2)	(3)	(4)	(5)	(6)
TWO	2	1	139	0.098	0.110	−0.027	0.092	0.105	0.081
				(0.098)	(0.101)	(0.152)	(0.199)	(0.097)	(0.288)
THREE	3	2	423	0.243***	0.243***	0.239***	0.230**	0.244***	0.232**
				(0.071)	(0.077)	(0.084)	(0.090)	(0.071)	(0.091)
THREE	3	1	613	0.301***	0.312***	0.281***	0.302***	0.304***	0.297***
				(0.067)	(0.070)	(0.078)	(0.086)	(0.067)	(0.084)
THREE	2	1	1306	0.241***	0.225***	0.258***	0.243***	0.242***	0.243***
				(0.035)	(0.035)	(0.039)	(0.039)	(0.035)	(0.091)
FOUR	4	3	15	0.636*				0.691	
				(0.370)				(0.479)	
FOUR	4	2	22	−0.181				−0.118	
				(0.290)				(0.282)	
FOUR	4	1	37	0.333	0.363	0.559		0.417	
				(0.300)	(0.256)	(1.189)		(0.322)	
FOUR	3	2	404	0.207***	0.195**	0.213**	0.177	0.207***	0.176
				(0.069)	(0.077)	(0.091)	(0.114)	(0.069)	(0.115)
FOUR	3	1	621	0.101*	0.094	0.130*	0.127	0.099*	0.126
				(0.058)	(0.060)	(0.077)	(0.082)	(0.058)	(0.082)
FOUR	2	1	1141	0.345***	0.340***	0.347***	0.339***	0.343***	0.337***
				(0.047)	(0.049)	(0.060)	(0.062)	(0.048)	(0.063)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 iterations) in parentheses. OLS regressions for equation (12) for mono-establishment firms. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Appendix D: Technological Intensity

Table A14: Technological Intensity: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	Low	Medium-Low	Medium-High	High
TWO	4,140	0.022	0.065***	0.053	−0.003
		(0.018)	(0.064)	(0.048)	(0.088)
THREE	18,114	0.257***	0.312***	0.467***	0.547***
		(0.020)	(0.028)	(0.034)	(0.069)
FOUR	14,633	0.392***	0.376***	0.496***	0.612***
		(0.025)	(0.035)	(0.035)	(0.077)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (10) with industry and location controls. The table only reports the value of the coefficient μ_1 from equation (10). The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker interacted with the degree of technological intensity of the industry he is employed in. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Table A15: Technological Intensity: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	Low	Medium-Low	Medium-High	High
THREE	3	2	419	0.325**	0.217	0.050	0.249
				(0.133)	(0.263)	(0.153)	(0.220)
THREE	3	1	602	0.371**	0.086	0.392**	0.584**
				(0.145)	(0.154)	(0.165)	(0.258)
THREE	2	1	1247	0.357***	0.199***	0.180*	0.219**
				(0.070)	(0.068)	(0.095)	(0.106)
FOUR	3	2	417	0.014	0.657**	0.378**	−0.032
				(0.119)	(0.281)	(0.175)	(0.202)
FOUR	3	1	617	0.197**	0.354**	0.145	0.008
				(0.094)	(0.179)	(0.158)	(0.263)
FOUR	2	1	1131	0.297***	0.445***	0.221**	0.290**
				(0.093)	(0.152)	(0.093)	(0.138)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 iterations) in parentheses. OLS regressions for equation (12) with industry and location controls. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$ interacted with the industry's degree of technological intensity. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Appendix D: Product Differentiation

Table A16: Product Differentiation: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	Low	High
TWO	4,414	-0.002 (0.019)	0.057*** (0.017)
THREE	19,751	0.270*** (0.024)	0.359*** (0.017)
FOUR	15,989	0.387*** (0.029)	0.444*** (0.019)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (10) with industry and location controls. The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker interacted with the degree of product differentiation of the industry he is employed in. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Table A17: Product Differentiation: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	Low	High
THREE	3	2	452	0.187 (0.129)	0.277** (0.109)
THREE	2	1	1374	0.207*** (0.066)	0.248*** (0.048)
FOUR	3	2	451	0.106 (0.132)	0.206 (0.140)
FOUR	3	1	686	0.131 (0.101)	0.117 (0.078)
FOUR	2	1	1248	0.132* (0.080)	0.371*** (0.068)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 iterations) in parentheses. OLS regressions for equation (12) with industry and location controls. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$ interacted with the industry's degree of product differentiation. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Appendix E: Robustness Checks Inconsistent Estimates

Table A18: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	(1)	(2)	(3)	(4)	(5)
TWO	1,746	0.118***	0.093***	0.110***	0.093***	0.254***
		(0.023)	(0.022)	(0.028)	(0.029)	(0.098)
THREE	10,374	0.634***	0.604***	0.611***	0.596***	0.805***
		(0.023)	(0.022)	(0.022)	(0.022)	(0.033)
FOUR	8,601	0.706***	0.685***	0.688***	0.677***	0.838***
		(0.028)	(0.028)	(0.028)	(0.028)	(0.036)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (10) for workers with at least 10 years in the dataset. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France.*

Table A19: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Size	(1)	(2)	(3)	(4)	(5)	(6)
TWO	2	1	38	0.354**	0.326	0.127		0.348**	
				(0.143)	(0.206)	(1.473)		(0.139)	
THREE	3	2	232	0.455***	0.449***	0.502***	0.513***	0.460***	0.511***
				(0.076)	(0.070)	(0.125)	(0.139)	(0.075)	(0.142)
THREE	3	1	380	0.421***	0.429***	0.284**	0.277**	0.412***	0.265**
				(0.092)	(0.092)	(0.123)	(0.133)	(0.092)	(0.135)
THREE	2	1	747	0.382***	0.373***	0.388***	0.373***	0.381***	0.370***
				(0.048)	(0.049)	(0.059)	(0.061)	(0.049)	(0.062)
FOUR	4	3	9	0.674				0.881	
				(0.551)				(1.000)	
FOUR	4	2	14	0.112				0.111	
				(0.204)				(0.246)	
FOUR	4	1	19	0.243	0.567	0.522		0.260	
				(0.591)	(1.089)	(2.524)		(0.645)	
FOUR	3	2	225	0.348***	0.358***	0.423***	0.391**	0.336***	0.369*
				(0.092)	(0.105)	(0.140)	(0.173)	(0.093)	(0.192)
FOUR	3	1	379	0.138***	0.138***	0.114	0.117	0.135**	0.127
				(0.070)	(0.071)	(0.109)	(0.124)	(0.072)	(0.129)
FOUR	2	1	679	0.357***	0.361***	0.318**	0.329**	0.356***	0.326**
				(0.081)	(0.083)	(0.124)	(0.131)	(0.080)	(0.132)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (12) for workers with at least 10 years. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Appendix E: Robustness Checks Estimation Error

Table A20: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	(1)	(2)	(3)	(4)	(5)
TWO	1,860	0.084***	0.064***	0.080***	0.058**	0.084
		(0.024)	(0.023)	(0.028)	(0.028)	(0.076)
THREE	7,374	0.528***	0.487***	0.489***	0.464***	0.490***
		(0.026)	(0.026)	(0.025)	(0.025)	(0.040)
FOUR	4,659	0.639***	0.602***	0.607***	0.585***	0.605***
		(0.035)	(0.036)	(0.036)	(0.038)	(0.052)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (10) for workers who have moved to a new firm in the year 2009. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The left-hand side variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Entries are omitted because the sample size was too small.*

Table A21: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	(1)	(2)	(3)	(4)	(5)	(6)
THREE	3	2	183	0.368***	0.290***	0.193	0.072	0.407***	0.078
				(0.083)	(0.087)	(0.135)	(0.121)	(0.086)	(0.147)
THREE	3	1	320	0.078	0.093	0.164	0.095	0.075	0.107
				(0.059)	(0.060)	(0.111)	(0.116)	(0.060)	(0.123)
THREE	2	1	439	0.127**	0.132***	0.072	0.099	0.089*	0.038
				(0.051)	(0.051)	(0.072)	(0.074)	(0.050)	(0.074)
FOUR	3	2	139	0.271**	0.212*	0.261	0.215	0.221**	0.165
				(0.110)	(0.113)	(0.166)	(0.260)	(0.114)	(0.243)
FOUR	3	1	237	0.100	0.094	0.080	0.121	0.083	0.072
				(0.066)	(0.071)	(0.101)	(0.114)	(0.064)	(0.127)
FOUR	2	1	318	0.304***	0.289***	0.234**	0.222*	0.291***	0.205
				(0.062)	(0.065)	(0.118)	(0.134)	(0.061)	(0.136)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Bootstrapped standard errors (500 replications) in parentheses. OLS regression results of equation (12) for workers who have moved to a new firm in 2009. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The left-hand side variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 17 industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*